The Multiple Hybrid Bootstrap
and
Frequency Domain Testing for Periodic Stationarity

Von der Carl-Friedrich-Gauß-Fakultät
Technische Universität Carolo-Wilhelmina zu Braunschweig
zur Erlangung des Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation

von	Carsten Jentsch
geboren am	27. Dezember 1979
in	Holzminden

Eingereicht am:	20. Oktober 2010
Mündliche Prüfung am:	17. Dezember 2010
Referent:	Prof. Dr. Jens-Peter Kreiß
Korreferent:	Prof. Dr. Efstathios Paparoditis
Korreferent:	Prof. Dr. Alexander Lindner

(2010)

To my parents

Contents

Chapter 1

Introduction

In general, time series analysis is concerned with statistical inference based on time series data, which is a sequence of data points measured at successive times. Typically, a realization of some length n of an unknown real-valued discrete stochastic process $(X_t, t \in \mathbb{Z})$ is observed, that is, one has data $X_1, \ldots, X_n$ at hand. Standard examples are stock prices, but also other measurements over time are conceivable and included in this general framework.

The purpose of time series analysis is to extract meaningful statistics and other characteristics from the data. In other words, the aim of time series analysis is to gather as much information of the underlying stochastic process $(X_t, t \in \mathbb{Z})$ that is contained in the observed data stretch as even possible. Unfortunately, the data generation process is absolutely unknown in most real life situations, which makes it very difficult, if not impossible to do statistical inference.

In classical time series analysis, the most popular assumption that is imposed on the underlying process $(X_t, t \in \mathbb{Z})$ to guarantee a certain amount of manageability is the assumption of stationarity. For instance, a strictly stationary process is a stochastic process whose joint probability distribution is invariant with respect to time shifts. In particular, this assumption implies that mean and autocovariance function are also shift-invariant when second moments of $(X_t, t \in \mathbb{Z})$ are supposed to exist. Precisely, it holds

$$E(X_t) = \mu, \tag{1.1}$$

$$Cov(X_{t+h}, X_t) = \gamma(h) \tag{1.2}$$

for all $t, h \in \mathbb{Z}$, that is, the mean is constant over time and the autocovariances depend only on the temporal distance, but not on the location. Stochastic processes that fulfill (1.1) and (1.2) are called weak, covariance or second order stationary. The very crucial property of second order stationarity makes it possible to base time series analysis on autocovariances $\gamma(h)$ or on autocorrelations $\rho(h)$, where $\rho(h) = \frac{\gamma(h)}{\gamma(0)}$, because these well-defined quantities can be estimated consistently from the data using their empirical counterparts

$$\widehat{\gamma}(h) = \begin{cases} \frac{1}{n}\sum_{t=1}^{n-|h|}(X_{t+|h|} - \overline{X})(X_t - \overline{X}), & |h| < n \\ 0, & \text{otherwise} \end{cases} \tag{1.3}$$

and $\widehat{\rho}(h) = \frac{\widehat{\gamma}(h)}{\widehat{\gamma}(0)}$, where $\overline{X} = \frac{1}{n}\sum_{t=1}^{n} X_t$ is the sample mean. Of course, second order stationarity allows also for meaningful investigation of the sample mean itself, because its limit in probability is now well-defined as well. Techniques that relate to autocovariances and autocorrelations to analyze the (linear) dependence structure of a stationary time series are often referred to as time-domain methods.

Another classical approach to time series analysis is the spectral analysis of the corresponding stochastic process. Corresponding techniques are usually called frequency-domain methods and deal in the majority of cases with the spectral density $f(\omega)$, which is a 2π-periodic non-negative function on the real line. This quantity is closely related to the autocovariance function $\gamma(h)$ defined above. Under its absolute summability, that is, $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$, it holds

$$f(\omega) = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty}\gamma(h)e^{-ih\omega}, \quad \omega \in \mathbb{R}. \tag{1.4}$$

If one is interested in estimation of the spectral density $f(\omega)$, it seems obvious to replace the unknown quantities $\gamma(h)$ by their estimates $\widehat{\gamma}(h)$ as defined in (1.3). This results in

$$\frac{1}{2\pi}\sum_{h=-(n-1)}^{n-1}\widehat{\gamma}(h)e^{-ih\omega} = \frac{1}{2\pi n}\left|\sum_{t=1}^{n}(X_t - \overline{X})e^{-it\omega}\right|^2 = I_{n,\overline{X}}(\omega),$$

where $I_{n,\overline{X}}(\omega)$ is called centralized periodogram. In the following chapters, it is usually assumed that $\mu = 0$, which allows the use of $I_n(\omega)$ with

$$I_n(\omega) = |J_n(\omega)|^2 = \frac{1}{2\pi n}\left|\sum_{t=1}^{n} X_t e^{-it\omega}\right|^2 \tag{1.5}$$

instead of $I_{n,\overline{X}}(\omega)$ above, where $J_n(\omega) = \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} X_t e^{-it\omega}$ is the discrete Fourier transform (DFT) of $X_1, \ldots, X_n$. Therefore, $I_n(\omega)$ is called periodogram from now on.

Unfortunately, the periodogram $I_n(\omega)$ is an asymptotically unbiased, but *not* consistent estimate of $f(\omega)$. Hence, kernel spectral density estimators $\widehat{f}(\omega)$ with

$$\widehat{f}(\omega) = \frac{1}{nh}\sum_{k=-\lfloor\frac{n-1}{2}\rfloor}^{\lfloor\frac{n}{2}\rfloor} K\left(\frac{\omega - \omega_k}{h}\right) I_n(\omega_k), \quad \omega \in \mathbb{R} \tag{1.6}$$

have to be used to get asymptotically consistent estimators of the spectral density function $f(\omega)$, where $\lfloor x \rfloor$ is the integer part of $x \in \mathbb{R}$, K is a nonnegative kernel function and h is the bandwidth.

It is worth noting that both approaches, time-domain and frequency-domain, have their own appeal and it depends heavily on the specific situation which appears to be more appropriate and has to be preferred.

However, the assumption of some sort of stationarity imposed on the stochastic process $(X_t, t \in \mathbb{Z})$ does usually not suffice to prove meaningful results as central limit theorems or to tackle the problem of reasonable forecasting in time series analysis. For this reason and to gain more structure that allows for asymptotic inference, it is often supposed that the underlying process is linear or that it fulfills some other kind of feasible dependence conditions as for instance mixing or weak dependence.

In this thesis, we are dealing with the assumption of linearity. A stochastic process $(X_t, t \in \mathbb{Z})$ is called linear time series, if it exhibits a representation

$$X_t = \sum_{\nu=-\infty}^{\infty} c_\nu e_{t-\nu}, \quad t \in \mathbb{Z}, \tag{1.7}$$

where $(c_\nu, \nu \in \mathbb{Z})$ with $c_0 = 1$ is an absolutely summable real-valued sequence and $(e_t, t \in \mathbb{Z})$ is a sequence of independent and identically distributed random variables with $E(e_t) = 0$ and $E(e_t^2) = \sigma_e^2 \in (0, \infty)$. Important special cases of (1.7) are causal moving-average processes of order q (MA(q) processes)

$$X_t = \sum_{\nu=0}^{q} c_\nu e_{t-\nu}, \quad t \in \mathbb{Z}, \tag{1.8}$$

but also autoregressive processes of order p (AR(p) processes)

$$X_t = \sum_{\nu=1}^{p} a_\nu X_{t-\nu} + e_t, \quad t \in \mathbb{Z} \tag{1.9}$$

and the combination of both, that is, autoregressive moving-average processes of orders p and q (ARMA(p,q) processes)

$$X_t - \sum_{\nu=1}^{p} a_\nu X_{t-\nu} = \sum_{\nu=0}^{q} c_\nu e_{t-\nu}, \quad t \in \mathbb{Z}, \tag{1.10}$$

are contained in the quite rich class of linear processes, where in all cases $p, q \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$.

Note that AR(p) and ARMA(p,q) processes require some additional assumptions on the roots of the corresponding polynomial $A(z) = 1 - \sum_{\nu=1}^{p} a_\nu z^\nu$ to be stationary and to be included in the class of linear processes.

In general, mathematical statistics is always concerned with uncertainty caused by randomness and so is time series analysis. As described above, quantities have to be estimated from the data and it is of canonical interest, whether the obtained estimates are close to their true, but unknown counterparts. Typically, to address this problem, asymptotic theory is necessary to construct $(1-\alpha)$-confidence intervals for some $\alpha \in (0,1)$. Under suitable assumptions on the underlying process $(X_t, t \in \mathbb{Z})$, it is often possible to derive a central limit theorem for the estimator of interest and its limiting distribution is then used to obtain the desired confidence interval of (approximate) level $1 - \alpha$.

Closely related to the construction of confidence intervals is the task of testing statistical hypothesis, where accurate critical values are needed to obtain tests that maintain some level α. Usually, approximate critical values are obtained also using asymptotic results on the limiting distributions of the test statistic.

In summary, both issues, construction of confidence intervals and derivation of critical values, are concerned with approximating unknown distributions of certain statistics of interest.

Although the time series process $(X_t, t \in \mathbb{Z})$ considered so far is supposed to be a one-dimensional real-valued process, of course, it is in many situations insufficient to deal with just one dimension. Hence, to give consideration to this issue, the processes dealt with in the following chapters are asumed to be of some arbitrary dimension, which may result in surprising and unexpected phenomena at times.

After this short introduction into the topic of time series analysis, its statistical purpose and some of the most important concepts, the upcoming Sections 1.1 and 1.2 deal briefly with the thematic environments of the main contents of this thesis. Afterwards, these are discussed elaborately in Chapter 2 and Chapter 3, respectively. For a more detailed introduction to time series analysis, we refer to Kreiss and Neuhaus (2006) and Brockwell and Davis (1991).

1.1 Bootstrapping dependent data

As already mentioned above, time series analysis has to deal with uncertainty caused by randomness in the observed data as every other subdiscipline of mathematical statistics. It is also emphasized that this task may be considered generally as the problem of approximating unknown distributions of statistics of interest, which is usually tackled with the help of asymptotic distributions obtained from central limit theorems. Primarily, this proceeding has two drawbacks. First of all, it is an asymptotic approach which causes the approximation to be reasonable just for large sample sizes. Second, the involved normal distribution forces the approximating distribution to be symmetric, even if the true distribution is heavily skewed. These facts possibly result in poor confidence intervals and critical values that are not reliable.

In the last decades, so-called resampling procedures have become widely accepted as tools to approximate unknown distributions. Typically, these techniques are less affected by the issues discussed above than the traditional approach that uses central limit theorems. Among others, the bootstrap scheme initially introduced by Efron (1979) for independent and identically distributed (i.i.d.) observations is one method that has found its way into the tool box of statisticians as well as practitioners.

In the situation of i.i.d. random variables $X_1, \ldots, X_n$, the idea of bootstrap can be briefly described as follows. To approximate the distribution of some statistic $T_n = T_n(X_1, \ldots, X_n)$, one draws randomly with replacement from the original observations n times to get a bootstrap data set $X_1^*, \ldots, X_n^*$, which is used to compute the corresponding statistic $T_n^* = T_n(X_1^*, \ldots, X_n^*)$. This procedure is executed B-times, where B is large and the empirical distribution of these B values of T_n^* is used to approximate the desired distribution.

The proceeding described above can only be reasoned with the i.i.d. assumption and the transfer of this idea to the case of dependent observations as encountered in time series analysis is not straightforward and therefore not trivial. In recent years various approaches for bootstrapping dependent data have been suggested in the literature to handle this task. Basically, one can distinguish between the following three types:

- Residual bootstrap methods (Kreiss (1988), Bühlmann (1997),...)
- Block bootstrap methods (Künsch (1989),...)
- Frequency domain methods (Franke and Härdle (1992), Dahlhaus and Janas (1996), Kreiss and Paparoditis (2003), Shao and Wu (2007)...)

Regarding the statistical properties of the three bootstrap types above, there are pros and cons either way and every type may appeal to the user because of its simplicity, universality or interpretability, for instance.

Residual bootstrap

In general, residual bootstrap methods have in common that some parametric (e.g. autoregressive) model is fitted to the data at first and the classical i.i.d. bootstrap is applied to the estimated residuals afterwards which are assumed to be i.i.d. random variables at least approximately. These techniques show usually quite good behaviour in simulation studies, but they are parametric approaches and, therefore, they work just for a restricted parametric class of time series models. Moreover, to approximate the dependence structure of the data sufficiently accurate, a very high order of the fitted model is often necessary, which results in a possible huge number of parameters that have to be estimated. This effect becomes even more problematic when dealing with multivariate time series data.

Block bootstrap

Block bootstrap methods are maybe the most straightforward generalization of the original scheme for i.i.d. data to the dependent case. Basically, they block the observed sample $X_1, \ldots, X_n$ in blocks of some length $l << n$ and one draws randomly with replacement from these blocks and glues them together to get a bootstrap data set $X_1^*, \ldots, X_n^*$. Under the assumption of strict stationarity, all blocks are identically distributed and drawing with replacement results in independency, which resembles the i.i.d. scheme as far as possible. It is worth noting, to obtain asymptotically consistent procedures, it is necessary to have block length and number of blocks tending to infinity with increasing sample size. These techniques are nonparametric and work under very general assumptions for this reason. But in situations, where certain parametric assumptions are actually satisfied, the block bootstrap performs usually considerably less accurate than its corresponding parametric counterpart that relies on approximately i.i.d. residuals.

Frequency domain bootstrap

In comparison to residual and block bootstrap methods that are both time-domain approaches, the literature on bootstrap schemes that apply in the frequency domain has increased substantially in recent years. Typically, they rely on asymptotic features of the periodogram that appear to be surprising at first sight. In fact, it is well-known that the periodogram ordinates evaluated at different frequencies $\omega, \lambda \in (0, \pi)$, $\omega \neq \lambda$ are asymptotically independent and that $I_n(\omega)$ is asymptotically exponentially distributed with parameter $f(\omega)$. These techniques do not require parametric assumptions and they usually show reasonable behaviour in simulations. However, the main handicap of these methods is that they resample the periodogram, but they do not have the ability to produce bootstrap replicates in the time domain, which restricts their applicability to functionals of the periodogram.

As discussed elaborately in Chapter 2 of this thesis, the intention of the hybrid bootstrap is to combine residual and frequency domain based methods to obtain a bootstrap proposal that satisfies the desired properties of both approaches. In particular, a bootstrap procedure is requested that is valid in a more general setup than the residual bootstrap, that delivers bootstrap replicates in the time domain and that shows good simulation results. So far, there is only little literature on bootstrap for multivariate time series, especially on nonparametric methods. Therefore, the case of vector-valued time series data is addressed particularly to fill this gap.

1.2 Stationarity vs. periodic stationarity

At the beginning of this chapter, it is emphasized that some kind of stationarity is usually necessary to make statistical inference possible. Any version of stationarity guarantees that certain distributional properties of the partially observed stochastic process $(X_t, t \in \mathbb{Z})$ do not change over time. For instance, second order stationarity allows for consistent estimation of autocovariances $\gamma(h)$, because more and more information on the linear dependence structure becomes available with increasing sample size. In other words, the key feature of stationarity is that all distributional properties under consideration of the stochastic process $(X_t, t \in \mathbb{Z})$ recur arbitrarily often when the sample size tends to infinity and that the time periods where they recur are known or follow at least some sufficiently known pattern.

Many non-stationary processes can be transformed into stationary or at least approximately stationary processes using de-trending or other techniques, but this is not always possible. For instance, in climatology and other geophysical sciences, one often encounters data from stochastic processes whose covariance structures appear to be non-stationary over time, but with a periodic behaviour. One may think about water-levels of a river measured monthly. Due to seasonal changing rainfall and other natural effects as annual snowmelt, it seems evident that autocovariances between February and May and between June and September, for example, do not coincide. But also economic time series as unemployment data comprise naturally periodic structures due to annually recurring effects that go beyond periodic trends.

Motivated by these considerations, the general notion of periodically correlated (or periodically stationary) processes as an extension of second order stationarity was introduced by Gladyshev (1961), who considered general random sequences without imposing any parametric assumptions. A few years later, Jones and Brelsford (1967) were the first to link the concept of periodic stationarity with the popular class of autoregressive processes. They assumed the coefficients in AR models [cf. (1.9)] not to be just constant, but to vary periodically with time to get a model that fits well into the concept of periodic stationarity.

The class of so-called periodic AR processes (PAR processes) and their properties have been investigated by many authors in the literature. Moreover, analogue generalizations of models (1.7), (1.8) and (1.10) have been suggested as well. Of course, these models are no longer stationary, but their distributional properties recur systematically enough to allow for meaningful statistical analysis. In fact, all these models share the crucial property that they may be represented as higher-dimensional stationary models. For instance, a d-variate PAR model with s periods can be written as an sd-variate AR model. This fundamental property makes it possible to use all well-established techniques for vector-valued stationary time series here also.

It is also worth noting that stationary models are contained in the more general class of periodically stationary models as special cases. Therefore, it is of canonical interest whether the data under consideration is actually generated by a truly periodically stationary model or just from a usual stationary model. This interest is reasoned by the fact that the number of involved parameters that have to be estimated from the data increases substantially when switching from stationary to periodically stationary models. Because of a possible s fold increase, where s is the number of periods, care must be exercised in its application. To hold the number of parameters in the model down, it is also important to choose the period properly. For instance, quarterly data is sometimes assumed to have period $s = 4$, but nevertheless, it is also imaginable that this kind of data has actual period $s = 2$.

Referring to the issue of a preferably parsimonious modelling of periodically stationary time series, a new test statistic is proposed in Chapter 3 that goes without any parametric assumptions on the underlying linear process. Essentially, the test is based on the nonparametric kernel estimate of a slightly adjusted spectral density matrix $\mathbf{g}(\omega)$ of the corresponding higher-dimensional stochastic process. Precisely, the suggested test exploits the specific shape of $\mathbf{g}(\omega)$ under the hypothesis of the underlying process to be stationary or periodically stationary with some smaller period.

Critical values for the testing procedure are obtain from the limiting distribution of a central limit theorem, but also by using the hybrid bootstrap scheme discussed in Chapter 2, which appears to be well suited for this purpose.

1.3 Acknowledgements

First of all, I would like to express my gratitude to my advisor Prof. Dr. Jens-Peter Kreiss for offering me a position in his group, introducing me into the subject and giving me the opportunity to prepare this thesis. In particular, I would like to thank him for the great confidence he placed all the time in me and my doing, whether research or teaching.

Also, I would like to thank Prof. Dr. Efstathios Paparoditis for supporting my research considerably during a research visit in Nicosia, Cyprus and for taking part in this doctoral examination procedure.

Chapter 2

The multiple hybrid bootstrap - Resampling multivariate linear processes

Based on: Carsten Jentsch and Jens-Peter Kreiss[1]
The multiple hybrid bootstrap - Resampling multivariate linear processes.
J. Multivariate Anal. **101** (2010), 2320 - 2345.

Abstract. The chapter reconsiders the autoregressive aided periodogram bootstrap (AAPB) which has been suggested in Kreiss and Paparoditis (2003). Their idea was to combine a time domain parametric and a frequency domain nonparametric bootstrap to mimic not only a part but as much as possible the complete covariance structure of the underlying time series. We extend the AAPB in two directions. Our procedure explicitly leads to bootstrap observations in the time domain and it is applicable to multivariate linear processes, but agrees exactly with the AAPB in the univariate case, when applied to functionals of the periodogram. The asymptotic theory developed shows validity of the multiple hybrid bootstrap procedure for the sample mean, kernel spectral density estimates and, with less generality, for autocovariances.

2000 *Mathematics Subject Classification.* 62G09, 62M10, 62H12.

Keywords and phrases. frequency domain bootstrap; multivariate bootstrap; multivariate linear time series; kernel estimators; discrete Fourier transform; Cholesky decomposition; spectral density matrix; autocovariance matrix; sample mean.

[1]Institut für Mathematische Stochastik, Technische Universität Braunschweig, Pockelsstraße 14, D-38106 Braunschweig, Germany, email: j.kreiss@tu-bs.de, http://www.tu-braunschweig.de/stochastik/

2.1 Introduction

In 1979, Efron's seminal paper [Efron (1979)] on the i.i.d. bootstrap as an extension of the jackknife initiated the fruitful theory of resampling methods in statistics. Since then a great many of papers concerning resampling techniques for i.i.d. as well as for non i.i.d. data has been proposed, whereas, by now, the i.i.d. case has been understood quite well. However, bootstrap methods have been acknowledged as a powerful tool for approximating certain distributional characteristics of statistics as, for example, variance or covariance, which are sometimes difficult to compute or even not possible to derive analytically. In particular, in time series analysis, due to the potentially complicated dependence structure of the data, often bootstrap methods are required to overcome this barrier, especially, if one wants to avoid the assumption of Gaussianity.

Besides parametric methods that are just applicable to a nonsatisfying narrow class of time series models, several nonparametric approaches for resampling dependent data have been suggested. For instance, Künsch (1989) introduced the so-called block-bootstrap, where blocks of data from a stationary process are resampled to preserve the dependence structure to some extent. See Bühlmann (2002), Lahiri (2003) and Härdle, Horowitz and Kreiss (2003) for an overview of existing methods.

In recent years, bootstrap procedures in the frequency domain have become more and more popular [compare Paparoditis (2002) for a survey]. Most of them are based on resampling the periodogram as in the paper by Franke and Härdle (1992), who proposed a nonparametric residual-based bootstrap that uses an initial (nonparametric) estimate of the spectral density and i.i.d. resampling of (appropriately defined) frequency domain residuals. They proved asymptotic validity for kernel spectral density estimates while Dahlhaus and Janas (1996) extended these validity to ratio statistics and Whittle estimators. Paparoditis and Politis (1999) followed an alternative approach exploiting smoothness properties of the spectral density and resample locally from adjacent periodogram ordinates. In an early unpublished manuscript, Hurvich and Zeger (1987) use the property that the relation between periodogram and spectral density can be described by means of a multiplicative regression model.

The idea of Kreiss and Paparoditis (2003) was to combine a time domain parametric and a frequency domain nonparametric bootstrap to widen the class of periodogram statistics for which their autoregressive aided periodogram bootstrap (AAPB) remains valid. They use a parametric (autoregressive) fit to catch the essential features of the data and to imitate the weak dependence structure of the periodogram ordinates while a nonparametric correction (in the frequency domain) is applied in order to catch features not represented by the parametric fit. Compare also Sergides and Paparoditis (2007), who carried over this idea to locally stationary processes.

However, the above mentioned frequency based resampling procedures share one handicap. All of them generate bootstrap periodogram replicates and, for this reason, can be applied to statistics that are functionals of the periodogram, exclusively.

In this chapter, we pick up the idea of the AAPB bootstrap introduced by Kreiss and Paparoditis (2003) and enhance their method in two directions. On the one hand, we modify the AAPB in such a manner that our new procedure has the ability to provide explicitly bootstrap replicates in the time domain. Further, we generalize our approach to the multivariate case, on the other hand. In doing so, we had to realize that indeed most of the univariate results are transferable one-to-one to the multivariate case, but also that this is *not* true in all situations.

Recently, Kirch and Politis (2009) proposed also a frequency-domain bootstrap scheme that is capable to generate time-domain replicates and is well suited for change point analysis.

So far, there is only little literature on bootstrap for multivariate time series, especially on nonparametric bootstrap methods. However, one dimension is evidently not enough to study the possibly sophisticated interdependencies between two or more quantities measured over time. Particularly, in econometric work, interest often centers on cross-variable dynamic interactions, which are frequently described with the concept of cointegration. For instance, in the case of a univariate linear time series, the empirical autocovariances concerning different lags obey a CLT with specific handsome covariance matrix in the limit [see Brockwell and Davis (1991), Proposition 7.3.1]. For this reason, using the Δ-method, the limiting covariance matrix of the empirical autocorrelations is not affected by the fourth order cumulant of the i.i.d. white noise process. This fact, in turn, allows the AAPB to be valid for autocorrelations and for ratio statistics in general. If one considers multivariate linear time series this does not remain true any longer. Compare Hannan (1970, Chapter IV, Section 3 and Theorem 14, p. 228) for the unattractive shape of the entrywise asymptotic covariance structure. Here, bootstrap methods may help approximating the distribution of these statistics.

Paparoditis (1996) considered a parametric bootstrap for vector-valued autoregressive time series of infinite order. The approach of Franke and Härdle (1992) has been extended to the multivariate case by Berkowitz and Diebold (1997) without proving validity. Dai and Guo (2004) proposed to smooth the Cholesky decomposition of a raw estimate of a multivariate spectrum, allowing different degrees of smoothness for different elements, while Guo and Dai (2006) extended their method to multivariate locally stationary processes. Goodness-of-fit tests for VARMA models are investigated by Paparoditis (2005), where the asymptotic distribution of the test statistic is established and therefore a bootstrap method is developed.

In the following we prove validity of our multiple hybrid bootstrap method under some mild general assumptions for the sample mean and for kernel spectral density estimators as well as asymptotic normality for empirical autocovariances, where the here proposed method is shown to work in some important special cases. Moreover, we check the validity for some statistics deduced from the above mentioned as, for example, cospectrum and quadrature spectrum.

In contrast to the AAPB paper, where all asymptotic results are derived for general classes of spectral means and ratio statistics, we restrict our considerations for the hybrid bootstrap in the multivariate setting to empirical autocovariances. Regarding their asymptotic behaviours in Theorem 2.5.3, it becomes clear that it is not possible to obtain validity for ratio statistics in general which would have been an analogue to Corollary 4.1 (ii) in Kreiss and Paparoditis (2003). However, a more general result corresponding to Theorem 4.1 (ii) in their paper for multivariate spectral means should be possible under suitable assumptions.

Also the case where the order of the autoregressive fit is allowed to tend to infinity with increasing sample size while assuming the underlying multivariate process to be causal and invertible is not considered here. This would correspond to Theorem 4.1 (i) and Corollary 4.1 (i) in the paper above, but analogue validity results are expected in the multivariate case as well.

This chapter is organized as follows. In Section 2.2, at first, we discuss our idea how to extend the AAPB to get bootstrap observations in the time domain and, thereafter, we generalize this concept to the multivariate case. The technical assumptions needed throughout the chapter are summarized in Section 2.3 while the multiple hybrid bootstrap procedure is described in detail in Section 2.4. Section 2.5 deals with applications of the suggested bootstrap in approximating the sampling behaviour of sample mean, spectral density estimates and empirical autocovariances as well as from these quantities deduced statistics. A small simulation study is presented in Section 2.6. Finally, proofs of the main results as well as of some technical lemmas are found in Section 2.7.

2.2 Preliminaries

We consider a strictly stationary r-dimensional process $\underline{\mathbf{X}} = (\underline{X}_t, t \in \mathbb{Z})$ and assume that $\underline{X}_t = (X_{t,1}, \ldots, X_{t,r})^T$ has the linear representation

$$\underline{X}_t = \sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu \underline{\epsilon}_{t-\nu}, \quad t \in \mathbb{Z}, \tag{2.1}$$

where $\mathbf{C}_\nu = (C_{\nu,ij})_{i,j=1,\ldots,r}$, $\nu \in \mathbb{Z}$ are $(r \times r)$ matrices, $\mathbf{C}_0 = \mathbf{I}_r$ is the $(r \times r)$ unit matrix and the sequence $(\mathbf{C}_\nu : \nu \in \mathbb{Z})$ is entrywise absolutely summable. Further, the error process $(\underline{\epsilon}_t, t \in \mathbb{Z})$ is assumed to consist of r-dimensional independent and identically distributed random variables $\underline{\epsilon}_t = (\epsilon_{t,1}, \ldots, \epsilon_{t,r})^T$ with $E[\underline{\epsilon}_t] = \underline{0}$ and $E[\underline{\epsilon}_t \underline{\epsilon}_t^T] = \mathbf{\Sigma}$, where the $(r \times r)$ covariance matrix $\mathbf{\Sigma}$ is supposed to be positive definite. Under these assumptions, $\underline{\mathbf{X}}$ exhibits the spectral density

$$\mathbf{f}(\omega) = \frac{1}{2\pi} \left(\sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu e^{-i\nu\omega} \right) \mathbf{\Sigma} \overline{\left(\sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu e^{-i\nu\omega} \right)}^T . \tag{2.2}$$

Here and in the following, we underline vector-valued quantities and write matrix-valued ones as bold letters. $\overline{\mathbf{Z}}$ denotes the (entrywise) complex conjugate of a matrix $\mathbf{Z}$ and X^T indicates the transpose of a vector or matrix X.

Since our first main intention is to pick up the concept of the AAPB bootstrap proposed by Kreiss and Paparoditis (2003) and modify it to obtain a procedure that is explicitly able to generate bootstrap replicates in the time domain, initially, we consider the univariate case $r = 1$ to simplify matters and sketch the steps of their method before demonstrating which step is the sticking point.

The univariate AAPB approach can be summarized as follows. With real-valued observations $X_1, \ldots, X_n$ at hand, Kreiss and Paparoditis apply a usual residual-based autoregressive bootstrap of fixed order $p \in \mathbb{N}$ to obtain bootstrap replicates $X_1^+, \ldots, X_n^+$ and compute the periodogram $I_n^+(\omega) = \frac{1}{2\pi n} | \sum_{t=1}^n X_t^+ e^{-it\omega}|^2$ of these quantities to switch over to the frequency domain. So far, this is just a parametric bootstrap that, of course, is not valid asymptotically if the underlying data does not stem from an autoregressive model of order less or equal to p. Therefore, they *correct* the periodogram $I_n^+(\omega)$ by multiplication with a nonparametric (pre-whitening) correction function $\widehat{q}(\omega)$, defined as

$$\widehat{q}(\omega) = \frac{1}{n} \sum_{j=-N}^{N} K_h(\omega - \omega_j) \frac{I_n(\omega_j)}{\widehat{f}_{AR}(\omega_j)}, \tag{2.3}$$

where $\omega_j = 2\pi \frac{j}{n}$, $N = [\frac{n}{2}]$, h is the bandwidth, K is a kernel function, $K_h(\cdot) = \frac{1}{h} K(\frac{\cdot}{h})$, $I_n(\omega)$ is the periodogram based on $X_1, \ldots, X_n$ and $\widehat{f}_{AR}$ is the spectral density obtained from the autoregressive fit. Their proceeding is motivated by the following facts. Recall that we want to bootstrap the periodogram $I_n(\omega)$ and under some assumptions on the coefficients of the linear representation of X_t in (2.1), it holds

$$E[I_n(\omega)] = f(\omega) + o(1), \tag{2.4}$$

but using the simple residual AR-bootstrap, however, yields

$$E^+[I_n^+(\omega)] = f_{AR}(\omega) + o_P(1), \tag{2.5}$$

where f is the true spectral density of the process $\mathbf{X}$ and f_{AR} is the spectral density of the theoretical autoregressive model of order p that is obtained as n tends to infinity. Note, that $f \neq f_{AR}$ in general. Here, as usual, E^+ denotes the conditional expectation given $X_1, \ldots, X_n$.

Since the estimate $\widehat{q}(\omega)$ in (2.3) converges to $\frac{f(\omega)}{f_{AR}(\omega)}$ in probability under some reasonable assumptions, their self-evident attempt to solve the problem argued in (2.4) and (2.5) is to design *corrected* bootstrap periodogram replicates $I_n^*(\omega)$ according to

$$I_n^*(\omega) = \widehat{q}(\omega) I_n^+(\omega),$$

obtaining

$$E^+[I_n^*(\omega)] = \widehat{q}(\omega) E^+[I_n^+(\omega)] = f(\omega) + o_P(1), \tag{2.6}$$

which, by now, agrees with the expectation in (2.4). Thus, the last equation emphasizes that, in a certain sense, the AAPB does the proper correction in the frequency domain. For this reason, one would expect this method to work for all statistics whose asymptotic distributional characteristics can be written in terms of the spectral density. But there are statistics with this property that cannot be written itself by means of the periodogram as, for instance, the sample mean. Recall that under some standard assumptions the following CLT holds true:

$$\mathcal{L}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_t\right) \Rightarrow \mathcal{N}(0, 2\pi f(0)). \tag{2.7}$$

However, using just the simple AR-bootstrap, under suitable assumptions, it holds

$$\mathcal{L}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} X_t^+ | X_1, \ldots, X_n\right) \Rightarrow \mathcal{N}(0, 2\pi f_{AR}(0)) \tag{2.8}$$

in probability. Considering solely (2.7) and (2.8), a naive idea to construct a bootstrap that works for the sample mean is to generate $X_1^+, \ldots, X_n^+$ and multiply the whole data set with $\sqrt{\widehat{q}(0)}$. Doing so, with Slutsky, we get

$$\begin{aligned} \mathcal{L}\left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sqrt{\widehat{q}(0)} X_t^+ | X_1, \ldots, X_n\right) \quad &\Rightarrow \quad \mathcal{N}(0, 2\pi f_{AR}(0) q(0)) \\ &= \quad \mathcal{N}(0, 2\pi f(0)) \end{aligned}$$

in probability, but this approach is just taylor-made for the sample mean and does not remain valid in other cases as spectral density estimation or for ratio statistics. Therefore, a different modification of the AAPB has to be developed to solve this problem, but we will come back to this issue later.

Now, to answer the question why the AAPB is not capable to deliver bootstrap replicates in the time domain, observe that $I_n^+(\omega)$, $\omega \in [-\pi, \pi]$ does not contain all the information that is contained in the data set $X_1^+, \ldots, X_n^+$. This means, on the one hand, computing the periodogram causes an irretrievable loss of information, but switching to the frequency domain is necessary to apply the nonparametric correction, on the other hand. To get rid of this inconvenience, note that for the periodogram at the Fourier frequencies $\omega_j = 2\pi \frac{j}{n}$, $j = 1, \ldots, n$, it holds

$$I_n^+(\omega_j) = |J_n^+(\omega_j)|^2 = J_n^+(\omega_j)\overline{J_n^+(\omega_j)},$$

where $J_n^+(\omega_j) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n X_t^+ e^{-it\omega_j}$ is the discrete Fourier transform (DFT) and there is a one-to-one correspondence between $X_1^+, \ldots, X_n^+$ and $J_n^+(\omega_1), \ldots, J_n^+(\omega_n)$.

These circumstances result in the idea to compute the DFT $J_n^+(\omega_1), \ldots, J_n^+(\omega_n)$ instead of the periodogram, multiply them with appropriate correction terms $\widetilde{q}(\omega_j)$ and use the ono-to-one correspondence to get back to the time domain. The canonical choice of the correction term is $\widetilde{q}(\omega) = \sqrt{\widehat{q}(\omega)}$ and to set

$$J_n^*(\omega_j) = \widetilde{q}(\omega_j) J_n^+(\omega_j), \quad j = 1, \ldots, n,$$

because with this definition, it holds

$$J_n^*(\omega_j)\overline{J_n^*(\omega_j)} = \widetilde{q}(\omega_j) J_n^+(\omega_j) \overline{\widetilde{q}(\omega_j) J_n^+(\omega_j)} = \widehat{q}(\omega_j) I_n^+(\omega_j) = I_n^*(\omega_j), \tag{2.9}$$

which is exactly the correction done in the AAPB method. Finally, we exploit the one-to-one correspondence of the DFT, to define bootstrap observations $X_1^*, \ldots, X_n^*$ via inverse DFT, that is,

$$X_t^* = \sqrt{\frac{2\pi}{n}} \sum_{j=1}^n J_n^*(\omega_j) e^{it\omega_j}, \quad t = 1, \ldots, n. \tag{2.10}$$

Now, that we have developed a bootstrap method that directly leads to bootstrap observations in the time domain and, moreover, contains the AAPB as a special case, let us consider the sample mean discussed in (2.7) and (2.8) again. Using the replicates defined in (2.10), we get

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{t=1}^n X_t^* &= \frac{1}{\sqrt{n}} \sqrt{\frac{2\pi}{n}} \sum_{j=1}^n J_n^*(\omega_j) \sum_{t=1}^n e^{it\omega_j} \\
&= \sqrt{2\pi} J_n^*(0) \\
&= \widetilde{q}(0) \frac{1}{\sqrt{n}} \sum_{t=1}^n X_t^+,
\end{aligned}$$

which is exactly the naive correction suggested earlier to construct a bootstrap that works for the sample mean, but contrary to the previous situation this new approach remains valid in all situations where the AAPB is already shown to work thanks to relation (2.9).

Taking everything into account, the above derived bootstrap constitutes a reasonable modification of the AAPB that is able to produce bootstrap replicates in the time domain and, for this reason, is applicable to a wider class of statistics. We call this proposal the (univariate) *hybrid bootstrap.*

Next, we generalize the hybrid bootstrap to the multivariate case. From now on, the data of interest is supposed to have some arbitrary dimension $r \geq 1$, but to appreciate the main difficulties adapting the univariate proposal derived above, consider the vector-valued case $r \geq 2$, only.

The first step of the hybrid bootstrap generalizes to a usual residual-based vector-autoregressive scheme to obtain $\underline{X}_1^+, \ldots, \underline{X}_n^+$. Further, the periodogram

$$\mathbf{I}_n^+(\omega_j) = \underline{J}_n^+(\omega_j)\overline{\underline{J}_n^+(\omega_j)}^T, \quad j = 1, \ldots, n$$

becomes a hermitian $(r \times r)$ matrix and the (multivariate) discrete Fourier transform (mDFT) $\underline{J}_n^+(\omega_j) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^n \underline{X}_t^+ e^{-it\omega_j}$ is now an r-dimensional column vector. Reconsidering (2.4) and (2.5) in the vector-valued case, it still holds

$$E[\mathbf{I}_n(\omega)] = \mathbf{f}(\omega) + o(1) \tag{2.11}$$

as well as

$$E^+[\mathbf{I}_n^+(\omega)] = \mathbf{f}_{AR}(\omega) + o_P(1), \tag{2.12}$$

with $\mathbf{I}_n(\omega)$, $\mathbf{f}(\omega)$ and $\mathbf{f}_{AR}(\omega)$ according to the univariate case.

Maintaining the property to produce bootstrap replicates in the time domain, consequently, we have to correct the mDFT. Now, this has to be done by multiplication with a suitable $(r \times r)$ matrix $\widetilde{\mathbf{Q}}(\omega_j)$, defining

$$\underline{J}_n^*(\omega_j) = \widetilde{\mathbf{Q}}(\omega_j)\underline{J}_n^+(\omega_j), \quad j = 1, \ldots, n.$$

Similar to the univariate equation (2.9), now, we get

$$\underline{J}_n^*(\omega_j)\overline{\underline{J}_n^*(\omega_j)}^T = \widetilde{\mathbf{Q}}(\omega_j)\underline{J}_n^+(\omega_j)\overline{\widetilde{\mathbf{Q}}(\omega_j)\underline{J}_n^+(\omega_j)}^T = \widetilde{\mathbf{Q}}(\omega_j)\mathbf{I}_n^+(\omega_j)\overline{\widetilde{\mathbf{Q}}(\omega_j)}^T. \tag{2.13}$$

Concerning (2.12), the last relation (2.13) asks for the correction term $\widetilde{\mathbf{Q}}(\omega)$ to converge in probability to its limit $\mathbf{Q}(\omega)$ [Observe the notation differing to the univariate case! For $r = 1$, it holds $\mathbf{Q}(\omega) = \sqrt{q(\omega)}$ instead of $\mathbf{Q}(\omega) = q(\omega)$.], which has to satisfy the equality

$$\mathbf{Q}(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T = \mathbf{f}(\omega) \tag{2.14}$$

to get the analogue result to equation (2.6) obtained in the univariate case, that is,

$$E^+[\mathbf{I}_n^*(\omega)] = \widetilde{\mathbf{Q}}(\omega)E^+[\mathbf{I}_n^+(\omega)]\overline{\widetilde{\mathbf{Q}}(\omega)}^T = \mathbf{f}(\omega) + o_P(1).$$

Now, to answer the question how $\widetilde{\mathbf{Q}}(\omega)$ has to be defined to achieve this property, suppose we knew that $\mathbf{f}(\omega)$ and $\mathbf{f}_{AR}(\omega)$ have some representations

$$\mathbf{f}(\omega) = \mathbf{G}(\omega)\overline{\mathbf{G}(\omega)}^T \quad \text{and} \quad \mathbf{f}_{AR}(\omega) = \mathbf{B}(\omega)\overline{\mathbf{B}(\omega)}^T. \tag{2.15}$$

Then, if the inverse of $\mathbf{B}(\omega)$ exists, it seems self-evident to set $\mathbf{Q}(\omega) = \mathbf{G}(\omega)\mathbf{B}^{-1}(\omega)$, obtaining

$$\mathbf{Q}(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T = \mathbf{G}(\omega)\mathbf{B}^{-1}(\omega)\mathbf{B}(\omega)\overline{\mathbf{B}(\omega)}^T\overline{\mathbf{B}^{-1}(\omega)}^T\overline{\mathbf{G}(\omega)}^T = \mathbf{f}(\omega),$$

and accordingly to construct a nonparametric estimator $\widetilde{\mathbf{Q}}(\omega)$ for this quantity $\mathbf{G}(\omega)\mathbf{B}^{-1}(\omega)$.

If $\mathbf{f}(\omega)$ and $\mathbf{f}_{AR}(\omega)$ are positive definite, their uniquely determined Cholesky decompositions as in (2.15) exist, where $\mathbf{G}(\omega)$ and $\mathbf{B}(\omega)$ have full rank. Thus, we can state $\widetilde{\mathbf{Q}}(\omega)$ in terms of estimates for $\mathbf{f}(\omega)$ and $\mathbf{f}_{AR}(\omega)$.

As in the univariate case, $\mathbf{f}(\omega)$ can be estimated nonparametrically by $\widehat{\mathbf{f}}(\omega)$ via smoothing the periodogram matrix and $\mathbf{f}_{AR}(\omega)$ is estimated by $\widehat{\mathbf{f}}_{AR}(\omega)$, which is obtained from the residual vector AR-bootstrap. Assuming $\mathbf{f}(\omega)$ to be positive definite, then, for sufficiently large sample size n in relation to r, the estimates $\widehat{\mathbf{f}}(\omega)$ and $\widehat{\mathbf{f}}_{AR}(\omega)$ are positive definite in probability. Hence, we can define

$$\widetilde{\mathbf{Q}}(\omega) = \widehat{\mathbf{G}}(\omega)\widehat{\mathbf{B}}^{-1}(\omega),$$

where $\widehat{\mathbf{f}}(\omega) = \widehat{\mathbf{G}}(\omega)\overline{\widehat{\mathbf{G}}(\omega)}^T$ and $\widehat{\mathbf{f}}_{AR}(\omega) = \widehat{\mathbf{B}}(\omega)\overline{\widehat{\mathbf{B}}(\omega)}^T$. Observe also the detailed illustration of this *multiple hybrid bootstrap* proposal in Section 2.4 and, in particularly, Remark 2.4.1 on the choice of $\widetilde{\mathbf{Q}}(\omega)$.

2.3 Assumptions

2.3.1 The data generation process

We assume the underlying process $\underline{\mathbf{X}}$ to satisfy the following assumptions:

(A1) $(\underline{X}_t, t \in \mathbb{Z})$ is a $\mathbb{R}^r$-valued linear strictly stationary process

$$\underline{X}_t = \sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu \underline{\varepsilon}_{t-\nu}, \quad t \in \mathbb{Z},$$

where $\mathbf{C}_\nu$, $\nu \in \mathbb{Z}$ are $(r \times r)$ coefficient matrices, $\mathbf{C}_0 = \mathbf{I}_r$ is the $(r \times r)$ unit matrix and for all $j, k = 1, \ldots, r$ the summability condition

$$\sum_{\nu=-\infty}^{\infty} |\nu||C_{\nu;j,k}| < \infty$$

holds true. Further, $\sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu z^\nu$ is supposed to be nonsingular on the unit circle, that is

$$det\left(\sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu z^\nu\right) \neq 0 \quad \forall z \in \mathbb{C} : |z| = 1.$$

(A2) The error process is assumed to be a *standard white noise* [compare Lütkepohl (2005), p.73], that means $(\underline{\epsilon}_t, t \in \mathbb{Z})$ constitutes a sequence of independent and identically distributed $\mathbb{R}^r$-valued random variables with $E[\underline{\epsilon}_t] = \underline{0}$ and $E[\underline{\epsilon}_t \underline{\epsilon}_t^T] = \mathbf{\Sigma}$, where the covariance matrix $\mathbf{\Sigma}$ is supposed to be positive definite. Further, for $i, j, k, l = 1, \ldots, r$ the expectation $E[\epsilon_{t,i}\epsilon_{t,j}\epsilon_{t,k}\epsilon_{t,l}] < \infty$ exists and $\kappa_4(i, j, k, l)$ denotes the fourth-order cumulant between $\epsilon_{t,i}, \epsilon_{t,j}, \epsilon_{t,k}$ and $\epsilon_{t,l}$.

(A3) The spectral density $\mathbf{f}$ in (2.2) of $\underline{\mathbf{X}}$ is (entrywise) three times continuously differentiable on $[-\pi, \pi]$ and accordingly on the real line, when understood as continuously extended.

2.3.2 The kernel function

(K1) K denotes a nonnegative kernel function with compact support $[-\pi, \pi]$. The Fourier transform k of K, that is,

$$k(u) = \int_{-\pi}^{\pi} K(x)e^{-ixu}dx,$$

is assumed to be a symmetric, continuous and bounded function satisfying $k(0) = 2\pi$. Hence, the kernel has the representation

$$K(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} k(u)e^{iux}du.$$

Note that $k(0) = 2\pi$ implies that $\frac{1}{2\pi}\int_{-\infty}^{\infty} K(u)du = 1$, while the symmetry of k implies the same property for K.

(K2) The Fourier transform k of K satisfies $\int_{-\infty}^{\infty} k^2(u)du < \infty$.

(K3) K is three times continuously differentiable on $[-\pi, \pi]$ and its derivatives fulfill the smoothness condition $K^{(d)}(-\pi) = K^{(d)}(\pi) = 0$ for all $d = 0, 1, 2$.

2.3.3 The bandwidth

(B1) $h = h(n) \to 0$ as $n \to \infty$ such that $nh \to \infty$.

(B2) $h = h(n) \to 0$ as $n \to \infty$ such that $(nh^4)^{-1} = O(1)$.

(B3) $h = h(n) \to 0$ as $n \to \infty$ such that $(nh^6)^{-1} = O(1)$.

2.4 The hybrid bootstrap procedure

In this section, first of all, we describe the multiple hybrid bootstrap motivated in Section 2.2 in detail and, afterwards, we give a couple of comments on the choice of the correction function $\widetilde{\mathbf{Q}}(\omega)$ and thereby arising difficulties. Moreover, we discuss the special case where no autoregressive model is fitted at all.

Step 1. Given the $\mathbb{R}^r$-valued observations $\underline{X}_1, \ldots, \underline{X}_n$, we fit a vector-autoregressive process of order $p \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}$ ($VAR(p)$-model). This leads to estimated coefficient matrices $\widehat{\mathbf{A}}_1(p), \ldots, \widehat{\mathbf{A}}_p(p)$ and covariance matrix $\widehat{\mathbf{\Sigma}}(p)$, which are obtained from the multivariate Yule-Walker equations. Consider the estimated residuals

$$\widehat{\underline{\epsilon}_t} = \underline{X}_t - \sum_{\nu=1}^{p} \widehat{\mathbf{A}}_\nu(p) \underline{X}_{t-\nu}, \quad t = p+1, \ldots, n$$

and denote $\widehat{F}_n^c$ the empirical distribution function of the standardized quantities

$$\widetilde{\underline{\epsilon}_t} = \widehat{\mathbf{L}}(p)^{-1} \left(\widehat{\underline{\epsilon}_t} - \frac{1}{n-p} \sum_{s=p+1}^{n} \widehat{\underline{\epsilon}_s} \right), \quad t = p+1, \ldots, n,$$

where

$$\widehat{\mathbf{L}}(p)\widehat{\mathbf{L}}(p)^T = \frac{1}{n-p} \sum_{t=p+1}^{n} \left(\widehat{\underline{\epsilon}_t} - \frac{1}{n-p} \sum_{s=p+1}^{n} \widehat{\underline{\epsilon}_s} \right) \left(\widehat{\underline{\epsilon}_t} - \frac{1}{n-p} \sum_{r=p+1}^{n} \widehat{\underline{\epsilon}_r} \right)^T$$

is the Cholesky decomposition of the covariance matrix of the centered residuals. That is, $\widehat{F}_n^c$ has mean $\underline{0}$ and the unit matrix $\mathbf{I}_r$ as covariance matrix.

Step 2. Generate bootstrap observations $\underline{X}_1^+, \ldots, \underline{X}_n^+$ according to the following vector autoregressive model of order p:

$$\underline{X}_t^+ = \sum_{\nu=1}^{p} \widehat{\mathbf{A}}_\nu(p) \underline{X}_{t-\nu}^+ + \widehat{\mathbf{\Sigma}}^{1/2}(p) \underline{\epsilon}_t^+,$$

where $(\underline{\epsilon}_t^+)$ is a sequence of i.i.d. random variables with cumulative distribution function $\widehat{F}_n^c$ (conditionally on the given observations $\underline{X}_1, \ldots, \underline{X}_n$) and $\widehat{\mathbf{\Sigma}}^{1/2}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)^T = \widehat{\mathbf{\Sigma}}(p)$ is the Cholesky decomposition. Now, the time series $(\underline{X}_t^+, t \in \mathbb{Z})$ has the spectral density

$$\widehat{\mathbf{f}}_{AR}(\omega) = \frac{1}{2\pi}\left(\mathbf{I}_r - \sum_{k=1}^{p} \widehat{\mathbf{A}}_k(p)e^{-ik\omega}\right)^{-1} \widehat{\mathbf{\Sigma}}(p) \left(\left(\overline{\mathbf{I}_r - \sum_{k=1}^{p} \widehat{\mathbf{A}}_k(p)e^{-ik\omega}}\right)^{-1}\right)^T.$$

Thereby, the used multivariate Yule-Walker estimates ensure that $\widehat{\mathbf{f}}_{AR}(\omega)$ is always well defined [cf. Whittle (1963)], that is

$$det\left(\mathbf{I}_r - \sum_{\nu=1}^{p} \widehat{\mathbf{A}}_\nu(p)z^\nu\right) \neq 0 \quad \forall z \in \mathbb{C} : |z| \leq 1.$$

Step 3. Compute the (multivariate) discrete Fourier transform (mDFT) of the bootstrap observations $\underline{X}_1^+, \ldots, \underline{X}_n^+$, that is

$$\underline{J}_n^+(\omega_j) = \frac{1}{\sqrt{2\pi n}} \sum_{t=1}^{n} \underline{X}_t^+ e^{-it\omega_j}, \quad j = 1, \ldots, n$$

at the Fourier frequencies $\omega_j = 2\pi\frac{j}{n}$, $j = 1, \ldots, n$. Notice, there is a one-to-one correspondence

$$\underline{X}_1^+, \ldots, \underline{X}_n^+ \quad \leftrightarrow \quad \underline{J}_n^+(\omega_1), \ldots, \underline{J}_n^+(\omega_n).$$

Step 4. Define the nonparametric correction function $\widetilde{\mathbf{Q}}(\omega) = \widehat{\mathbf{G}}(\omega)\widehat{\mathbf{B}}(\omega)^{-1}$, where $\widehat{\mathbf{G}}(\omega)$ and $\widehat{\mathbf{B}}(\omega)$ are obtained via the following Cholesky decompositions (in lower triangular matrix times its transposed complex conjugate):

$$\widehat{\mathbf{B}}(\omega)\overline{\widehat{\mathbf{B}}(\omega)}^T = \widehat{\mathbf{f}}_{AR}(\omega) \tag{2.16}$$

and

$$\begin{aligned} &\widehat{\mathbf{G}}(\omega)\overline{\widehat{\mathbf{G}}(\omega)}^T \\ = \ &\widehat{\mathbf{B}}(\omega)\left(\frac{1}{n}\sum_{k=-N}^{N} K_h(\omega-\omega_k)\widehat{\mathbf{B}}(\omega_k)^{-1}\mathbf{I}_n(\omega_k)\overline{\widehat{\mathbf{B}}(\omega_k)^{-1}}^T\right)\overline{\widehat{\mathbf{B}}(\omega)}^T, \end{aligned} \tag{2.17}$$

whereas $N = [\frac{n}{2}]$, K is a kernel function, $K_h(\cdot) = \frac{1}{h}K(\frac{\cdot}{h})$ and h is the bandwidth. Furthermore, $\mathbf{I}_n(\omega) = \underline{J}_n(\omega)\overline{\underline{J}_n(\omega)}^T$ denotes the periodogram matrix of the given observations with

$$\underline{J}_n(\omega) = \frac{1}{\sqrt{2\pi n}}\sum_{t=1}^{n} \underline{X}_t e^{-it\omega}.$$

Now, compute the nonparametric estimator $\widetilde{\mathbf{Q}}$ at the Fourier frequencies $\omega_j = 2\pi\frac{j}{n}$, $j = 1, \ldots, n$. In doing so, all involved quantities are understood as periodically extended to the real line.

Step 5. At first, compute the mDFT $\underline{J}_n^+(\omega_j)$, $j = 1, \ldots, n$ of the parametrically via residual bootstrap generated observations $\underline{X}_1^+, \ldots, \underline{X}_n^+$ and afterwards apply the nonparametric correction function $\widetilde{\mathbf{Q}}(\omega)$ to get the *corrected* version of the mDFT, that is

$$\underline{J}_n^*(\omega_j) = \widetilde{\mathbf{Q}}(\omega_j)\underline{J}_n^+(\omega_j), \quad j = 1, \ldots, n.$$

Step 6. According to the inverse mDFT, the bootstrap observations $\underline{X}_1^*, \ldots, \underline{X}_n^*$ are defined as follows:

$$\underline{X}_t^* = \sqrt{\frac{2\pi}{n}} \sum_{j=1}^n \underline{J}_n^*(\omega_j) e^{it\omega_j}, \quad t = 1, \ldots, n.$$

Remark 2.4.1 (On the choice of $\widetilde{\mathbf{Q}}(\omega)$)**.**

(i) As illustrated in (2.15)*, basically, it is possible to use alternative decompositions e.g. square-root or Cholesky-decomposition in upper triangular matrix times its transposed complex conjugate. Although Cholesky needs positive definiteness, we choose this decomposition (in lower triangular matrix times its transposed complex conjugate), because it is uniquely defined and it automatically generates invertible matrices.*

(ii) Moreover, regarding just (2.14) *and* (2.15)*, it would even work if one uses different decompositions in* (2.16) *and* (2.17)*. This would lead to the same results in Section 2.5 except for Corollary 2.5.4, which will not remain valid, anymore.*

(iii) In definition (2.17)*, we follow the advice of Kreiss and Paparoditis (2003) and define* $\widehat{\mathbf{G}}(\omega)$ *via a nonparametric pre-whitening estimate of* $\mathbf{f}(\omega)$*. Asymptotically, we get the same results if we just set*

$$\widetilde{\mathbf{G}}(\omega)\overline{\widetilde{\mathbf{G}}(\omega)}^T = \frac{1}{n}\sum_{k=-N}^{N} K_h(\omega - \omega_k)\mathbf{I}_n(\omega_k)$$

and redefine $\mathring{\mathbf{Q}}(\omega) = \widetilde{\mathbf{G}}(\omega)\widehat{\mathbf{B}}(\omega)^{-1}$*, but for small sample sizes we expect slightly better results using* $\widehat{\mathbf{G}}(\omega)$*. Note, in the univariate case,* $\widetilde{\mathbf{Q}}(\omega)$ *agrees with* $\widetilde{q}(\omega)$ *as defined previous to* (2.9)*.*

(iv) Assumption (A1) *guarantees the positive definiteness of* $\mathbf{f}$ *and, for this reason, the pre-whitening estimate in* (2.17) *satisfies this property asymptotically (in*

probability). However, for very small sample sizes n relative to the dimension r, it may happen that the quantities on the right-hand sides of (2.16) *and* (2.17) *are just positive semidefinite and not positive definite, which in turn disallows computation of their Cholesky decompositions. For medium and large sample sizes this problem practically does not occur. Hence, it is advisable to define*

$$\widetilde{\mathbf{Q}}(\omega) = \begin{cases} \widehat{\mathbf{G}}(\omega)\widehat{\mathbf{B}}(\omega)^{-1}, & \widehat{\mathbf{G}}(\omega) \text{ and } \widehat{\mathbf{B}}(\omega) \text{ exist} \\ \mathbf{I}_r, & \text{otherwise} \end{cases}$$

to overcome this difficulty of well-definition. Observe that in the second case, the hybrid bootstrap becomes the usual residual AR-bootstrap.

(v) To obtain $\widetilde{\mathbf{Q}}(\omega)$ that satisfies (2.14) *in its limit (in probability), it is essential to estimate $\mathbf{f}_{AR}(\omega)$ and $\mathbf{f}(\omega)$ separately and decompose them first, before defining $\widetilde{\mathbf{Q}}(\omega)$ as its product.*

Remark 2.4.2 (On the choice of p)**.**
A common bootstrap technique in time series analysis is the autoregressive residual bootstrap, but often a high order p has to be chosen to capture the dependence structure properly. Regarding multiple time series data, this may result in a huge number of parameters to be estimated. Elaborate simulation studies done using the multiple hybrid bootstrap have shown very reasonable results even in the case $p = 1$. To point out the effect of the nonparametric correction and to underline the quality of the obtained bootstrap results, we choose $p = 1$ in our simulation study in Section 2.6, only.

Remark 2.4.3 (The special case $p = 0$)**.**
Setting $p = 0$ means that we do not fit any autoregressive model to the data $\underline{X}_1, \ldots, \underline{X}_n$ at all in Step 1 of our proposal. Actually, Step 2 shrivels to the standard i.i.d. bootstrap scheme obtaining $\underline{X}_1^+, \ldots, \underline{X}_n^+$. Although this ignores completely the dependence structure in $\underline{X}_1, \ldots, \underline{X}_n$, nevertheless, the hybrid bootstrap remains valid as discusssed later in Section 2.5. In comparison, the nonparametric residual-based periodogram bootstrap (NPB) proposed by Franke and Härdle (1992) uses that the periodogram ordinates are asymptotically independently distributed according to an exponential distribution. For this reason, they resample in the frequency domain to obtain i.i.d. exponentially distributed random variates. In the case $p = 0$, in contrast, we do i.i.d. resampling in the time domain disregarding the dependence in the data and switch to the frequency domain afterwards by computing the discrete Fourier transform. Observe that periodogram ordinates are just asymptotically independent, but for finite n this is not true anymore. Although we ignore the dependence contained in $\underline{X}_1, \ldots, \underline{X}_n$ by using this i.i.d. scheme setting $p = 0$, in comparison to the NPB, we get correlated periodogram ordinates in the frequency domain.

2.5 Asymptotic theory and validity

This section is organized in three subsections. In the first one, we state the validity of our procedure for the multivariate sample mean, which constitutes an extension of the AAPB introduced by Kreiss and Paparoditis (2003), also in the univariate case. Validity for kernel spectral density matrix estimation and related quantities is discussed in the second subsection and, finally, the third deals with the asymptotic covariance structure of (entries of) empirical autocovariance matrices, their weak convergence in general as well as validity in some special situations. In the following, we use repeatedly Mallows' d_2-metric [cf. Mallows (1972)]. The d_2-distance between distributions $\mathcal{P}_1$ and $\mathcal{P}_2$ is defined as follows:

$$d_2\{\mathcal{P}_1, \mathcal{P}_2\} = \inf\{E|Y_1 - Y_2|^2\}^{1/2},$$

where the infimum is taken over all joint distributions for the pair of random variables Y_1 and Y_2 whose fixed marginal distributions are $\mathcal{P}_1$ and $\mathcal{P}_2$, respectively. Compare Bickel and Freedman (1981) for a detailed discussion and related results.

2.5.1 Sample mean

Theorem 2.5.1 (Validity for the sample mean).
Suppose the assumptions (A1), (A2), (K1) and (B1) are satisfied. Then for all fixed $p \in \mathbb{N}_0$*, it holds*

$$d_2\left\{\mathcal{L}(\sqrt{n}\ \overline{\underline{X}}), \mathcal{L}(\sqrt{n}\ \overline{\underline{X}}^* | \underline{X}_1, \ldots, \underline{X}_n)\right\} \to 0$$

in probability, where $\overline{\underline{X}} = \frac{1}{n}\sum_{t=1}^{n} \underline{X}_t$ *and* $\overline{\underline{X}}^* = \frac{1}{n}\sum_{t=1}^{n} \underline{X}_t^*$.

2.5.2 Spectral density estimates

Theorem 2.5.2 (Validity for spectral density estimates).
Suppose the assumptions (A1), (A2), (A3), (K1), (K2), (K3) and (B3) are satisfied as well as $nb^5 \to C^2$ *with a constant* $C \geq 0$*. Then for all fixed orders* $p \in \mathbb{N}_0$ *of the autoregressive fit, all* $s \in \mathbb{N}$ *and arbitrary frequencies* $\omega_1, \ldots, \omega_s$ *(not necessarily Fourier frequencies), it holds*

$$\begin{aligned} d_2\Big\{&\mathcal{L}(\sqrt{nb}(\widehat{f}_{jk}(\omega_l) - f_{jk}(\omega_l)) : j,k = 1,\ldots,r; l = 1,\ldots,s), \\ &\mathcal{L}(\sqrt{nb}(\widehat{f}^*_{jk}(\omega_l) - \widetilde{f}_{jk}(\omega_l))|\underline{X}_1,\ldots,\underline{X}_n : j,k = 1,\ldots,r; l = 1,\ldots,s)\Big\} \to 0 \end{aligned}$$

in probability, where $\widehat{\mathbf{f}}(\omega) = \frac{1}{n}\sum_{j=-N}^{N} K_b(\omega - \omega_j)\mathbf{I}_n(\omega_j)$*,* $\widehat{\mathbf{f}}^*(\omega) = \frac{1}{n}\sum_{j=-N}^{N} K_b(\omega - \omega_j)\mathbf{I}_n^*(\omega_j)$ *and* $\widetilde{\mathbf{f}}(\omega) = \widetilde{\mathbf{Q}}(\omega)\widehat{\mathbf{f}}_{AR}(\omega)\overline{\widetilde{\mathbf{Q}}(\omega)}^T$.

A direct consequence of the above Theorem 2.5.2 is the corresponding result for the so-called cospectrum and quadrature spectrum, which are real-valued quantities and for this reason sometimes preferred to the complex-valued cross-spectral densities.

Corollary 2.5.1 (Cospectrum and quadrature spectrum).
Putting $\mathbf{f}(\omega) = \frac{1}{2}(\mathbf{c}_{spec}(\omega) - i\mathbf{q}_{spec}(\omega))$ *(analogue for* $\widehat{\mathbf{f}}(\omega)$, $\widehat{\mathbf{f}}^*(\omega)$ *and* $\widetilde{\mathbf{f}}(\omega)$*), we call the (real) matrix-valued quantities* $\mathbf{c}_{spec}(\omega)$ *and* $\mathbf{q}_{spec}(\omega)$ *the co- and quadrature spectral density matrices. Under the assumptions of Theorem 2.5.2 the following holds:*

$$\begin{aligned}
&d_2 \Big\{ \mathcal{L}(\sqrt{nb}(\widehat{c}_{spec,jk}(\omega_l) - c_{spec,jk}(\omega_l)) : j,k = 1,\ldots,r; l = 1,\ldots,s), \\
&\quad \mathcal{L}(\sqrt{nb}(\widehat{c}^*_{spec,jk}(\omega_l) - \widetilde{c}_{spec,jk}(\omega_l)) | \underline{X}_1,\ldots,\underline{X}_n : j,k = 1,\ldots,r; l = 1,\ldots,s) \Big\} \to 0, \\
&d_2 \Big\{ \mathcal{L}(\sqrt{nb}(\widehat{q}_{spec,jk}(\omega_l) - q_{spec,jk}(\omega_l)) : j,k = 1,\ldots,r; l = 1,\ldots,s), \\
&\quad \mathcal{L}(\sqrt{nb}(\widehat{q}^*_{spec,jk}(\omega_l) - \widetilde{q}_{spec,jk}(\omega_l)) | \underline{X}_1,\ldots,\underline{X}_n : j,k = 1,\ldots,r; l = 1,\ldots,s) \Big\} \to 0
\end{aligned}$$

in probability, respectively.

2.5.3 Empirical autocovariances

Autocovariances provide a lot of information about the stochastic dependence properties of a multivariate time series $\underline{\mathbf{X}}$. For instance, if one is interested in construction of confidence intervals, especially in the multivariate case, it is difficult to use existing central limit theorems to derive confidence regions. This is up to the sophisticated covariance matrix of the asymptotic normal distribution. Defining

$$\widehat{\mathbf{\Gamma}}(h) = \begin{cases} \frac{1}{n}\sum_{t=1}^{n-h}(\underline{X}_{t+h} - \overline{\underline{X}})(\underline{X}_t - \overline{\underline{X}})^T, & h \geq 0 \\ \frac{1}{n}\sum_{t=1-h}^{n}(\underline{X}_{t+h} - \overline{\underline{X}})(\underline{X}_t - \overline{\underline{X}})^T, & h < 0 \end{cases}, \tag{2.18}$$

namely, it holds [compare Hannan (1970), Chapter IV, Section 3 and Theorem 14, p. 228]

$$\begin{aligned}
& nCov(\widehat{\gamma}_{jk}(g) - \gamma_{jk}(g), \widehat{\gamma}_{lm}(h) - \gamma_{lm}(h)) \\
\to & \sum_{s_1,s_2,s_3,s_4=1}^{r} \left(\sum_{\nu_1=-\infty}^{\infty} C_{\nu_1,js_1} C_{\nu_1-g,ks_2} \right) \kappa_4(s_1,s_2,s_3,s_4) \left(\sum_{\nu_2=-\infty}^{\infty} C_{\nu_2,ls_3} C_{\nu_2-h,ms_4} \right) \\
& + \sum_{t=-\infty}^{\infty} \gamma_{km}(t)\gamma_{jl}(t-h+g) + \sum_{t=-\infty}^{\infty} \gamma_{kl}(t-h)\gamma_{jm}(t+g) \qquad (2.19) \\
= & \sum_{s_1,s_2,s_3,s_4=1}^{r} \left(\int_{-\pi}^{\pi} (\mathbf{C}(\omega_1))_{js_1} \left(\overline{\mathbf{C}(\omega_1)}^T \right)_{s_2 k} e^{ig\omega_1} d\omega_1 \right) \\
& \quad \times \kappa_4(s_1,s_2,s_3,s_4) \left(\int_{-\pi}^{\pi} (\mathbf{C}(\omega_2))_{ls_3} \left(\overline{\mathbf{C}(\omega_2)}^T \right)_{s_4 m} e^{ih\omega_2} d\omega_2 \right) \qquad (2.20) \\
& + \sum_{t=-\infty}^{\infty} \gamma_{km}(t)\gamma_{jl}(t-h+g) + \sum_{t=-\infty}^{\infty} \gamma_{kl}(t-h)\gamma_{jm}(t+g)
\end{aligned}$$

for all $j,k,l,m = 1,\ldots,r$ and all lags $g,h \in \mathbb{Z}$, where $\mathbf{C}(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu e^{-i\nu\omega}$ is the transfer function of $\underline{\mathbf{X}}$ and $\kappa_4(s_1,s_2,s_3,s_4)$ is the fourth order joint cumulant between $\epsilon_{t,s_1}, \epsilon_{t,s_2}, \epsilon_{t,s_3}$ and ϵ_{t,s_4}.

The first sums in (2.19) and (2.20) containing these cumulants are difficult to handle and to interpret. For this reason, bootstrap methods may possibly help to overcome this difficulty. Desirable is to have a bootstrap procedure that is able to replicate the covariance structure above as far as possible.

In the following two Theorems 2.5.3 and 2.5.4, we state the asymptotics for the hybrid bootstrap corresponding to (2.20) on the bootstrap level.

Theorem 2.5.3 (Asymptotic covariance structure).
Assume that (A1), (A2), (K1) and (B2) are satisfied and let $p \in \mathbb{N}_0$. Defining $\mathbf{C}_p(\omega) = \frac{1}{\sqrt{2\pi}} \sum_{\nu=0}^{\infty} \mathbf{C}_\nu(p) e^{-i\nu\omega}$, where $\mathbf{C}_\nu(p)$, $\nu \in \mathbb{N}_0$ are the coefficient matrices of the causal representation of the best autoregressive fit of order p to $\underline{\mathbf{X}}$ in L_2-distance, for all $j,k,l,m = 1,\ldots,r$ and all $g,h \in \mathbb{Z}$, the following convergence in probability holds true:

$$
\begin{aligned}
& nCov^+(\widehat{\gamma}^*_{jk}(g) - E^+[\widehat{\gamma}^*_{jk}(g)], \widehat{\gamma}^*_{lm}(h) - E^+[\widehat{\gamma}^*_{lm}(h)]) \\
\to \quad & \sum_{s_1,s_2,s_3,s_4=1}^{r} \left(\int_{-\pi}^{\pi} (\mathbf{Q}(\omega_1)\mathbf{C}_p(\omega_1))_{js_1} \left(\overline{\mathbf{C}_p(\omega_1)}^T \overline{\mathbf{Q}(\omega_1)}^T \right)_{s_2k} e^{ig\omega_1} d\omega_1 \right) \\
& \times \kappa_4(p; s_1,s_2,s_3,s_4) \left(\int_{-\pi}^{\pi} (\mathbf{Q}(\omega_2)\mathbf{C}_p(\omega_2))_{ls_3} \left(\overline{\mathbf{C}_p(\omega_2)}^T \overline{\mathbf{Q}(\omega_2)}^T \right)_{s_4m} e^{ih\omega_2} d\omega_2 \right) \\
& + \sum_{t=-\infty}^{\infty} \gamma_{km}(t)\gamma_{jl}(t-h+g) + \sum_{t=-\infty}^{\infty} \gamma_{kl}(t-h)\gamma_{jm}(t+g), \qquad (2.21)
\end{aligned}
$$

where Cov^+ is the conditional covariance given $\underline{X}_1,\ldots,\underline{X}_n$, $\widehat{\mathbf{\Gamma}}^*(h)$ is the bootstrap analogue of (2.18) *and $\kappa_4(p;s_1,s_2,s_3,s_4)$ is the fourth order joint cumulant between the corresponding components of the (non-standardized) residuals obtained by the best autoregressive fit.*

Theorem 2.5.4 (Asymptotic normality).
Suppose the assumptions (A1), (A2), (A3), (K1), (K3) and (B3) are satisfied. Then for all fixed $p \in \mathbb{N}_0$, all $s \in \mathbb{N}_0$ and lags $l = 0,\ldots,s$, it holds

$$\mathcal{L}(\sqrt{n}(\widehat{\gamma}^*_{jk}(l) - E^+[\widehat{\gamma}^*_{jk}(l)])|\underline{X}_1,\ldots,\underline{X}_n : j,k = 1,\ldots,r; l = 0,\ldots,s) \Rightarrow \mathcal{N}(\underline{0}, \mathbf{V})$$

in probability. Here, the asymptotic covariance matrix $\mathbf{V}$ can be constructed by the results of Theorem 2.5.3.

Unfortunately, the multiple hybrid bootstrap method does not work completely satisfactory in the general setting for autocovariances. In comparison to the AAPB this is not surprising, because in the univariate case the AAPB is just able to mimic the asymptotic distribution for autocorrelations (ratio statistics) and not for autocovariances (spectral means), where the arising fourth order cumulant of the white noise process is not captured properly [compare Theorem 4.1 (ii) and Corollary 4.1 (ii) in Kreiss and Paparoditis (2003)]. However, under suitable assumptions, a more general result for multivariate spectral means corresponding to Theorem 4.1 (i) in their paper is expected. The following direct corollary shows that our bootstrap procedure provides the same results as the AAPB in the univariate case for empirical autocovariances.

Corollary 2.5.2 (Univariate case).
Let $r = 1$. Under the assumptions of Theorem 2.5.4 we get

$$\mathcal{L}(\sqrt{n}(\widehat{\gamma}^*(l) - E^+[\widehat{\gamma}^*(l)])|X_1, \ldots, X_n : l = 0, \ldots, s) \Rightarrow \mathcal{N}(0, V),$$

where V is obtained by

$$\begin{aligned} & nCov^+(\widehat{\gamma}(g), \widehat{\gamma}(h)) \\ \to \quad & \gamma(g)\gamma(h)(\eta(p) - 3) + \sum_{t=-\infty}^{\infty} \gamma(t)\gamma(t-h+g) + \sum_{t=-\infty}^{\infty} \gamma(t-h)\gamma(t+g) \end{aligned} \tag{2.22}$$

in probability, where we set $E[(X_p - \sum_{\nu=1}^{p} a_\nu(p)X_{p-\nu})^2] = \sigma^2(p)$ and $E[(X_p - \sum_{\nu=1}^{p} a_\nu(p)X_{p-\nu})^4] = \eta(p)\sigma^4(p)$ and $a_\nu(p)$, $\nu = 1, \ldots, p$ are the coefficients of the best autoregressive fit of order p in L_2-distance.

Comparing (2.21) and (2.22) one striking difference regarding the first summands becomes obvious. The asymptotic covariance in the univariate case discussed in Corollary 2.5.2 depends exclusively through $\eta(p)$ which is related to the fourth order cumulant $\kappa_4(p)$ on the initially fitted autoregressive model and therefore on the underlying hybrid bootstrap proposal. In contrast, the complicated covariance structure derived in Theorem 2.5.3 depends on the fourth order joint cumulants $\kappa_4(p; s_1, s_2, s_3, s_4)$ and, additionally, on the correction function $\mathbf{Q}(\omega)$ as well as on the transfer function $\mathbf{C}_p(\omega)$ of the best autoregressive fit. The reason why these quantities do not vanish asymptotically for $r \geq 2$ is given in the following remark.

Remark 2.5.1.
The nonparametric correction achieved by $\mathbf{Q}(\omega)$ works properly only in the case when the multiplication is executed on either side of the spectral density matrix $\mathbf{f}_{AR}(\omega)$, that is,

$$\mathbf{Q}(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T = \mathbf{Q}(\omega)\mathbf{C}_p(\omega)\mathbf{\Sigma}(p)\overline{\mathbf{C}_p(\omega)}^T\overline{\mathbf{Q}(\omega)}^T = \mathbf{f}(\omega),$$

but one-sided application of $\mathbf{Q}(\omega)$ to the transfer function $\mathbf{C}_p(\omega)$ yields

$$\mathbf{Q}(\omega)\mathbf{C}_p(\omega) \neq \mathbf{C}(\omega)\mathbf{\Sigma}^{1/2}\mathbf{\Sigma}(p)^{-1/2} \tag{2.23}$$

in general. Observe that equality in (2.23) *is necessary for the quantities* $\mathbf{Q}(\omega)$ *and* $\mathbf{C}_p(\omega)$ *to disappear in* (2.21) *and, consequently, for the integrals to collapse in* (2.21) *obtaining a representation similar to* (2.19)*. This problem does not arise in the univariate case, where the square root of a positive real number is uniquely determined up to its sign, which is not true for the generalized square root of a positive definite matrix.*

Note that all these quantities $\kappa_4(p; s_1, s_2, s_3, s_4)$, $\mathbf{Q}(\omega)$ and $\mathbf{C}_p(\omega)$ depend on the order p of the autoregressive fit, which in turn causes the hybrid bootstrap as well as the AAPB to be not valid in general for empirical autocovariances. Due to this specific feature in the multivariate situation, moreover, it is neither possible to obtain validity for empirical autocorrelations under general assumptions nor for a more general class of ratio statistics. However, compared to the usual residual AR-bootstrap, the hybrid bootstrap is at least able to mimic exactly the second and third term in (2.19). In the upcoming Corollaries 2.5.3 and 2.5.4, two important special cases are presented where the hybrid bootstrap still works asymptotically.

Apparently, both methods (AAPB and hybrid bootstrap) do not have the ability to imitate the fourth moments and accordingly the fourth order cumulants of the underlying white noise process $(\underline{\epsilon}_t, t \in \mathbb{Z})$ properly. This problem does not appear if we assume a normal distribution for the error process, because in this case all occurring fourth order cumulants vanish and we immediately obtain the following result.

Corollary 2.5.3 (Gaussian case)**.**
Assume that the residuals $(\underline{\epsilon}_t, t \in \mathbb{Z})$ *are multivariate normally distributed. Under the assumptions of Theorem 2.5.4 for all* $s \in \mathbb{N}_0$ *and lags* $l = 0, \ldots, s$*, it holds*

$$d_2 \big\{ \mathcal{L}(\sqrt{n}(\widehat{\gamma}_{jk}(l) - E[\widehat{\gamma}_{jk}(l)]) : j,k = 1,\ldots,r; l = 0,\ldots,s), \\ \mathcal{L}(\sqrt{n}(\widehat{\gamma}^*_{jk}(l) - E^+[\widehat{\gamma}^*_{jk}(l)]) | \underline{X}_1,\ldots,\underline{X}_n : j,k = 1,\ldots,r; l = 0,\ldots,s) \big\} \to 0$$

in probability.

Assuming the underlying process $\underline{\mathbf{X}}$ to be a causal vector autoregressive time series of finite order $p_0 \in \mathbb{N}_0$ is another very important case. In this situation the usual residual bootstrap works well if we fit a model of order $p \geq p_0$. For this reason, we do not want the correction function $\widetilde{\mathbf{Q}}(\omega)$ to adjust anything and expect the hybrid bootstrap to be valid particularly in this case. Otherwise, this would represent a significant drawback compared to the residual bootstrap. The forthcoming corollary reinforces our speculation.

Corollary 2.5.4 ($VAR(p_0)$ case)**.**
Assume that the underlying observations $\underline{X}_1, \ldots, \underline{X}_n$ *originate from a causal* $VAR(p_0)$ *model with* $p_0 \in \mathbb{N}_0$*, that is, the stationary process* $\underline{\mathbf{X}}$ *satisfies*

$$\underline{X}_t = \sum_{k=1}^{p_0} \mathbf{A}_k \underline{X}_{t-k} + \underline{\epsilon}_t, \quad t \in \mathbb{Z}.$$

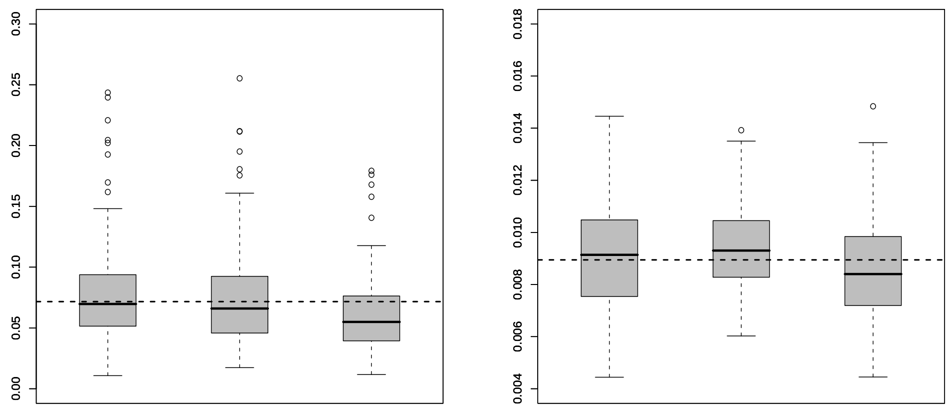

Figure 2.1: Boxplots of the bootstrap distributions for the variance of the first component of the sample mean in the $VMA(1)$ case with target indicated by the horizontal dashed line. In both panels from left to right: hybrid bootstrap (HB), AR-bootstrap (ARB) and moving block bootstrap (MBB). Left panel: $n = 50$, HB with $p = 1$ and $h = 0.3$; ARB with $p = 1$; MBB with $l = 5$. Right panel: $n = 400$, HB with $p = 1$ and $h = 0.15$; ARB with $p = 1$; MBB with $l = 10$.

Under the assumptions of Theorem 2.5.4 for all $p \in \mathbb{N}_0$, $p \geq p_0$, *all* $s \in \mathbb{N}_0$ *and lags* $l = 0, \ldots, s$, *it holds*

$$d_2 \left\{ \mathcal{L}(\sqrt{n}(\widehat{\gamma}_{jk}(l) - E[\widehat{\gamma}_{jk}(l)]) : j,k = 1,\ldots,r; l = 0,\ldots,s), \right.$$
$$\left. \mathcal{L}(\sqrt{n}(\widehat{\gamma}^*_{jk}(l) - E^+[\widehat{\gamma}^*_{jk}(l)]) | \underline{X}_1, \ldots, \underline{X}_n : j,k = 1,\ldots,r; l = 0,\ldots,s) \right\} \to 0$$

in probability.

Using techniques similar to those employed by Kreiss and Paparoditis (2003) proving their Theorem 4.1 (i), it seems also possible to achieve validity for empirical autocovariances (and for spectral means and ratio statistics in general) in the case of an underlying causal $VAR(\infty)$ model allowing the order $p = p(n)$ of the autoregressive fit to increase at an appropriate rate with the sample size n without assuming Gaussianity. Basically, this is because the correction term $\widetilde{\mathbf{Q}}(\omega)$ tends to the unit matrix in this case as well.

2.6 A simulation study

In this section we compare the performance of the proposed multiple hybrid bootstrap to that of the usual autoregressive bootstrap and that of the moving block

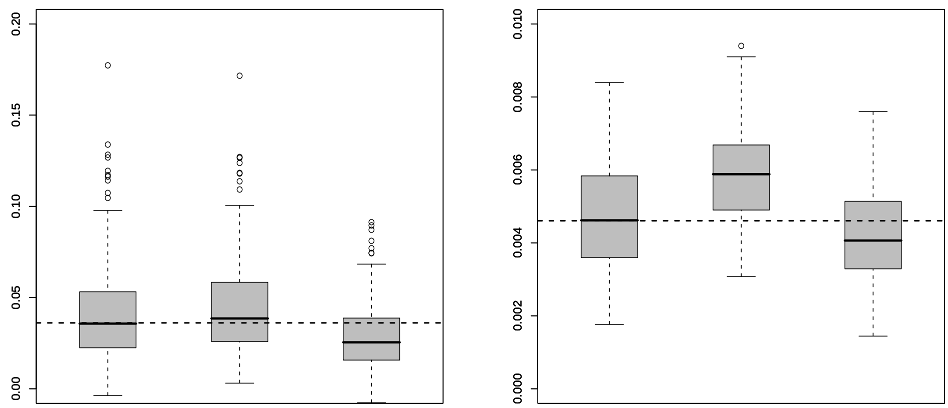

Figure 2.2: Boxplots of the bootstrap distributions for the covariance of both components of the sample mean in the $VMA(1)$ case with target indicated by the horizontal dashed line. In both panels from left to right: hybrid bootstrap (HB), AR-bootstrap (ARB) and moving block bootstrap (MBB). Left panel: $n = 50$, HB with $p = 1$ and $h = 0.3$; ARB with $p = 1$; MBB with $l = 5$. Right panel: $n = 400$, HB with $p = 1$ and $h = 0.15$; ARB with $p = 1$; MBB with $l = 10$.

bootstrap by means of simulation. In order to make such a comparison, we have chosen statistics for which all methods lead to asymptotically correct approximations. In particular, we study and compare the performance of the aforementioned bootstrap methods in estimating a) the variance σ^2 of the first component and b) the covariance γ_{12} of both components of the sample mean $\overline{\underline{X}} = \frac{1}{n}\sum_{t=1}^{n} \underline{X}_t$ of a bivariate time series data set.

Realizations of length $n = 50$ and $n = 400$ from two models

$$\underline{X}_t = \mathbf{A}_1 \underline{\epsilon}_{t-1} + \underline{\epsilon}_t \qquad \text{and} \qquad \underline{X}_t = \mathbf{A}_1 \underline{X}_{t-1} + \underline{\epsilon}_t$$

with i.i.d. $\underline{\epsilon}_t \sim \mathcal{N}(\underline{0}, \mathbf{\Sigma})$ have been considered, where the first one is a vector moving average model of order one ($VMA(1)$ model) and the second is a vector autoregressive model of order one ($VAR(1)$ model). In both cases, we have used

$$\mathbf{A}_1 = \begin{pmatrix} 0.5 & 0.9 \\ 0.0 & 0.5 \end{pmatrix}, \quad \mathbf{\Sigma} = \begin{pmatrix} 1.0 & 0.2 \\ 0.2 & 1.0 \end{pmatrix}.$$

To estimate the exact variance σ^2 and covariance γ_{12}, $10,000$ Monte-Carlo replications have been used while the bootstrap approximations are based on $B = 300$ bootstrap replications and we have simulated $M = 200$ data sets, respectively. In all

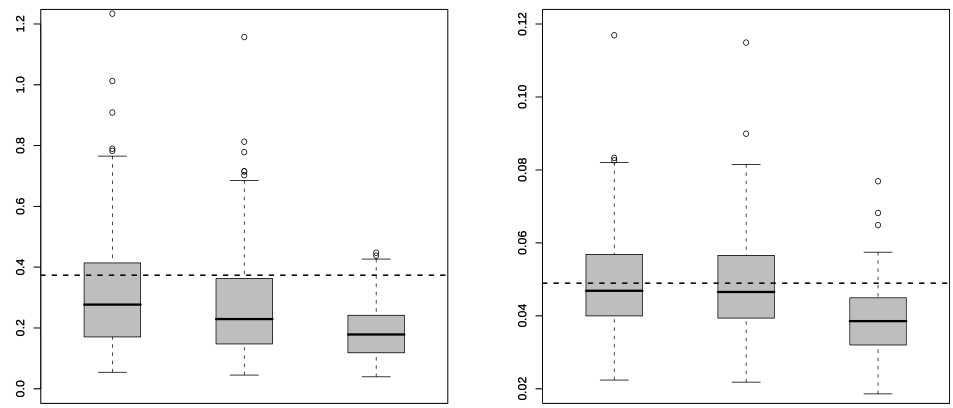

Figure 2.3: Boxplots of the bootstrap distributions for the variance of the first component of the sample mean in the $VAR(1)$ case with target indicated by the horizontal dashed line. In both panels from left to right: hybrid bootstrap (HB), AR-bootstrap (ARB) and moving block bootstrap (MBB). Left panel: $n = 50$, HB with $p = 1$ and $h = 0.3$; ARB with $p = 1$; MBB with $l = 5$. Right panel: $n = 400$, HB with $p = 1$ and $h = 0.15$; ARB with $p = 1$; MBB with $l = 10$.

cases, the Bartlett-Priestley kernel K has been used and an autoregressive model of order $p = 1$ is fitted to the data. Compare also Remark 2.4.2 concerning the choice of $p = 1$.

In Figures 2.1 - 2.4, some boxplots of the distributions of the different bootstrap approximations for the cases $n = 50$ and $n = 400$ are presented. To check how sensitive the hybrid bootstrap reacts concerning the choice of the bandwidth h in Figure 2.5 and 2.6 boxplots with different bandwidths are shown.

All figures show reasonable results for the hybrid bootstrap in comparison to the other methods, but the effect of the nonparametric correction is clearly seen in Figure 2.2, where the bias of the pure autoregressive bootstrap is reduced significantly. Moreover, as expected, the hybrid bootstrap works well for autoregressive time series data as illustrated in Figure 2.3 and 2.4, where even some bias reduction can be seen in comparison to the autoregressive bootstrap. The Figures 2.5 and 2.6 demonstrate that the hybrid bootstrap seems not to be over sensitive concerning the choice of h. In particular, the right panel in Figure 2.5 shows the typical behaviour of decreasing fluctuation with increasing bandwidth.

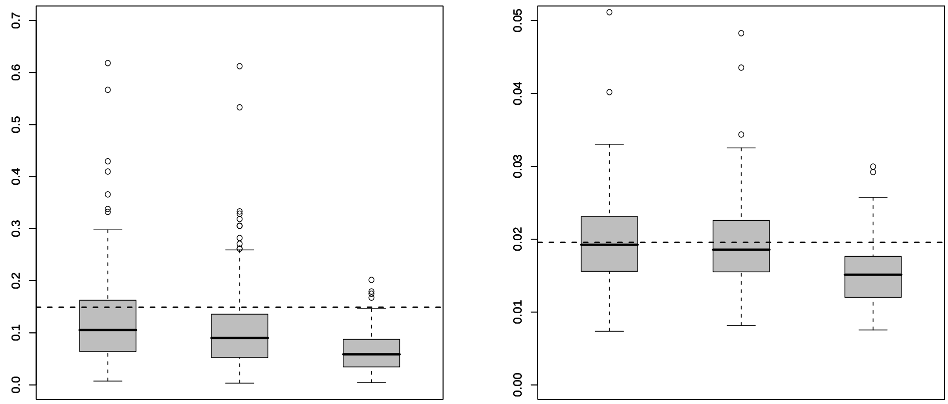

Figure 2.4: Boxplots of the bootstrap distributions for the covariance of both components of the sample mean in the $VAR(1)$ case with target indicated by the horizontal dashed line. In both panels from left to right: hybrid bootstrap (HB), AR-bootstrap (ARB) and moving block bootstrap (MBB). Left panel: $n = 50$, HB with $p = 1$ and $h = 0.3$; ARB with $p = 1$; MBB with $l = 5$. Right panel: $n = 400$, HB with $p = 1$ and $h = 0.15$; ARB with $p = 1$; MBB with $l = 10$.

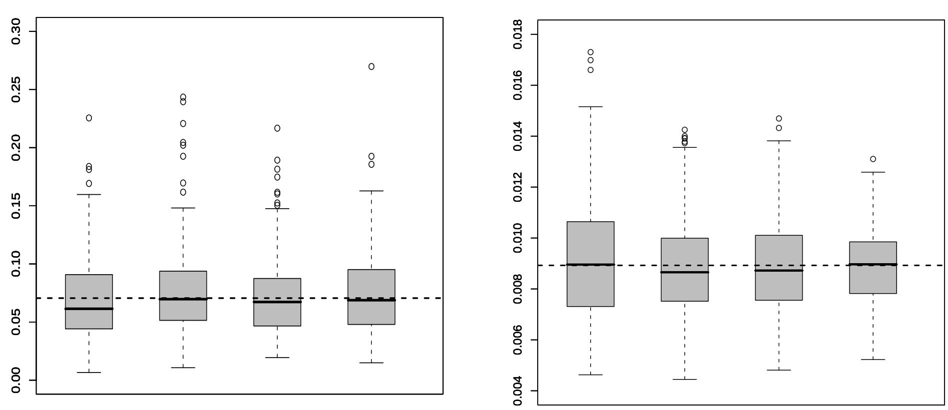

Figure 2.5: Boxplots of the bootstrap distributions for the variance of the first component of the sample mean using hybrid bootstrap (HB) in the $VMA(1)$ case with target indicated by the horizontal dashed line for different bandwidths h. Left panel: $n = 50$, from left to right: $h = 0.2$, $h = 0.3$, $h = 0.4$ and $h = 0.5$. Right panel: $n = 400$, from left to right: $h = 0.1$, $h = 0.15$, $h = 0.2$ and $h = 0.25$.

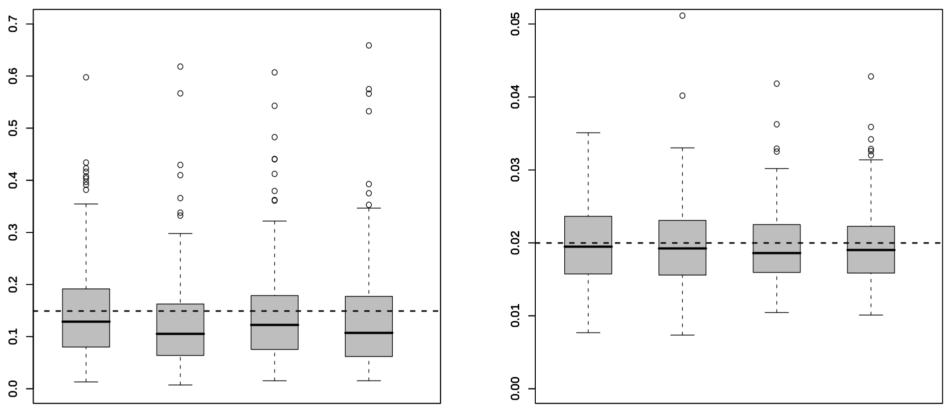

Figure 2.6: Boxplots of the bootstrap distributions for the covariance of both components of the sample mean using hybrid bootstrap (HB) in the $VAR(1)$ case with target indicated by the horizontal dashed line for different bandwidths h. Left panel: $n = 50$, from left to right: $h = 0.2$, $h = 0.3$, $h = 0.4$ and $h = 0.5$. Right panel: $n = 400$, from left to right: $h = 0.1$, $h = 0.15$, $h = 0.2$ and $h = 0.25$.

2.7 Proofs and auxiliary results

2.7.1 The nonparametric correction function

Lemma 2.7.1 (Consistency of the correction function).
Assume (A1), (A2), (K1) and (B1). Then, for the nonparametric correction function $\widetilde{\mathbf{Q}}(\omega) = \widehat{\mathbf{G}}(\omega)\widehat{\mathbf{B}}(\omega)^{-1}$ *as defined in* (2.16) *and* (2.17) *(note the suppressed dependence on the sample size n), it holds*

$$\widetilde{\mathbf{Q}}(\omega) \to \mathbf{Q}(\omega)$$

in probability for all ω, where $\mathbf{Q}(\omega) = \mathbf{G}(\omega)\mathbf{B}(\omega)^{-1}$ *with Cholesky decompositions* $\mathbf{B}(\omega)\overline{\mathbf{B}(\omega)}^T = \mathbf{f}_{AR}(\omega)$ *and* $\mathbf{G}(\omega)\overline{\mathbf{G}(\omega)}^T = \mathbf{f}(\omega)$*. If even (B2) is satisfied, we get the uniform convergence*

$$\sup_\omega \|\widetilde{\mathbf{Q}}(\omega) - \mathbf{Q}(\omega)\| = o_P(1)$$

and if additionally (A3), (K3) and (B3) are fulfilled, the first three (entrywise) derivatives of $\widetilde{\mathbf{Q}}(\omega)$ *exists and we get the uniform convergence in probability of the first two, that is*

$$\sup_\omega \|\widetilde{\mathbf{Q}}^{(j)}(\omega) - \mathbf{Q}^{(j)}(\omega)\| = o_P(1), \quad j = 1, 2$$

and the boundedness in probability of the third, that is, $\sup_{\omega}\|\widetilde{\mathbf{Q}}^{(3)}(\omega)\| = O_P(1)$.

Proof.
First of all, we discuss some preliminary considerations. The Cholesky decomposition $\mathbf{B}\overline{\mathbf{B}}^T = \mathbf{A}$ of a (complex) positive definite matrix $\mathbf{A}$ is obtained recursively by

$$b_{kl} = \begin{cases} 0, & k < l \\ (a_{kk} - \sum_{j=1}^{k-1} b_{kj}\overline{b_{kj}})^{1/2}, & k = l \\ \frac{1}{b_{ll}}(a_{kl} - \sum_{j=1}^{l-1} b_{kj}\overline{b_{lj}}), & k > l \end{cases}, \tag{2.24}$$

where $\mathbf{B}$ is uniquely defined and all diagonal elements are real-valued and strictly positive and therefore $\mathbf{B}$ is invertible. Assuming a matrix-valued function $\mathbf{A}(\omega)$ to be positive definite for all ω, the same properties hold for its Cholesky decomposition $\mathbf{B}(\omega)$. Further, if we assume $\mathbf{A}(\omega)$ to be (entrywise) k-times differentiable in ω, this property is also satisfied for $\mathbf{B}(\omega)$, which can be seen easily computing the derivatives according to (2.24). Moreover, if $(\mathbf{A}_n(\omega) : n \in \mathbb{N})$ is a sequence of matrix-valued functions assumed to be positive definite as well as k-times (entrywise) differentiable for all ω, uniform convergence of their first k derivatives $\mathbf{A}_n^{(d)}(\omega)$, $d = 0, 1, \ldots, k$ causes uniform convergence of the k-th derivative $\mathbf{B}_n^{(k)}(\omega)$ of the corresponding Cholesky decomposition $\mathbf{B}_n(\omega)$.

Since the spectral densities $\mathbf{f}(\omega)$ and $\mathbf{f}_{AR}(\omega)$ are both positive definite for all ω due to the assumptions (A1) and (A2) and because the Yule-Walker estimates always yield to stable autoregressive models [compare Whittle (1963)], it suffices to restrict considerations to the convergence of the quantities on the right-hand sides of (2.16) and (2.17) to $\mathbf{f}(\omega)$ and $\mathbf{f}_{AR}(\omega)$ respectively as well as the convergence of their derivatives. We prove only the most sophisticated assertion for $\widetilde{\mathbf{Q}}^{(2)}(\omega)$.

The uniform convergence of $\widehat{\mathbf{f}}_{AR}(\omega)$ in probability follows by standard arguments using (2.30) below and, because of the positive definiteness of its limit $\mathbf{f}_{AR}(\omega)$, we can treat $\widehat{\mathbf{f}}_{AR}(\omega)$ as a positive definite matrix for sufficiently large n (in probability). Hence, the right-hand side in (2.17) is well defined for large n (in probability). Entrywise geometrically decaying coefficient matrices of the causal representation of the (stable) autoregressive fit yield uniform convergence for all derivatives of $\widehat{\mathbf{f}}_{AR}(\omega)$ and the same holds true for its inverse $\widehat{\mathbf{f}}_{AR}^{-1}(\omega)$, causing the k-th derivatives of $\widehat{\mathbf{B}}(\omega)$ and $\widehat{\mathbf{B}}^{-1}(\omega)$ to converge uniformly, also. Now, consider the term on the right-hand side of (2.17) more closely and define

$$\widehat{\mathbf{Q}}(\omega) = \frac{1}{n}\sum_{k=-N}^{N} K_h(\omega - \omega_k)\widehat{\mathbf{B}}(\omega_k)^{-1}\mathbf{I}_n(\omega_k)\overline{\widehat{\mathbf{B}}(\omega_k)^{-1}}^T.$$

Thanks to the uniform convergence of $\widehat{\mathbf{B}}(\omega)$ and $\widehat{\mathbf{B}}^{-1}(\omega)$ and their derivatives, it remains to show

$$\sup_\omega \|\widehat{\mathbf{Q}}^{(d)}(\omega) - (\mathbf{B}^{-1}(\omega)\mathbf{f}(\omega)\overline{\mathbf{B}^{-1}(\omega)}^T)^{(d)}\| = o_P(1).$$

A Taylor series expansion yields

$$\begin{aligned} \widehat{\mathbf{Q}}^{(d)}(\omega) &= \sum_{s_1,s_2=0}^{d} (\widehat{\mathbf{B}}^{-1}(\omega))^{(s_1)} \left(\frac{1}{nh^{d+1}} \sum_{k=-N}^{N} K^{(d)}\left(\frac{\omega-\omega_k}{h}\right)(\omega_k-\omega)^{s_1+s_2}\mathbf{I}_n(\omega_k)\right) \\ &\quad \times\overline{(\widehat{\mathbf{B}}^{-1}(\omega))^{(s_2)}}^T + O_P(h) \end{aligned} \tag{2.25}$$

uniformly in ω and it remains to check the following uniform convergence for the expression in the big round parentheses in (2.25):

$$\sup_\omega \|\frac{1}{nh^{d+1}} \sum_{k=-N}^{N} K^{(d)}\left(\frac{\omega-\omega_k}{h}\right)(\omega_k-\omega)^{s}\mathbf{I}_n(\omega_k) - \frac{d!}{(d-s)!}\mathbf{f}^{(d-s)}(\omega)\| = o_P(1)$$

for $d = 0, 1, 2$ and $s = 0, 1, \ldots, d$. Observe that all sums in (2.25) with $s = s_1 + s_2 > d$ can be neglected because they vanish asymptotically with $O_P(h^{s-d})$ due to assumption (K3). To prove the last assertion, we follow the idea of Franke and Härdle (1992, Theorem A1). Initially, the last supremum is bounded by

$$\sup_\omega \|\frac{1}{nh^{d+1}} \sum_{k=-N}^{N} K^{(d)}\left(\frac{\omega-\omega_k}{h}\right)(\omega_k-\omega)^{s} \tag{2.26}$$

$$\times(\mathbf{I}_n(\omega_k) - \mathbf{C}(\omega_k)\mathbf{I}_{n,\epsilon}(\omega_k)\overline{\mathbf{C}(\omega_k)}^T)\|$$

$$+ \sup_\omega \|\frac{1}{nh^{d+1}} \sum_{k=-N}^{N} K^{(d)}\left(\frac{\omega-\omega_k}{h}\right)(\omega_k-\omega)^{s}\mathbf{C}(\omega_k)\mathbf{I}_{n,\epsilon}(\omega_k)\overline{\mathbf{C}(\omega_k)}^T \tag{2.27}$$

$$-\frac{\mathbf{f}^{(d-s)}(\omega)}{(d-s)!}\frac{1}{nh} \sum_{k=-N}^{N} K^{(d)}\left(\frac{\omega-\omega_k}{h}\right)\left(\frac{\omega_k-\omega}{h}\right)^d \|$$

$$+ \sup_\omega \|\frac{\mathbf{f}^{(d-s)}(\omega)}{(d-s)!}\frac{1}{nh} \sum_{k=-N}^{N} K^{(d)}\left(\frac{\omega-\omega_k}{h}\right)\left(\frac{\omega_k-\omega}{h}\right)^d - \frac{d!\mathbf{f}^{(d-s)}(\omega)}{(d-s)!}\|, \tag{2.28}$$

where $\mathbf{C}(\omega) = \sum_{\nu=-\infty}^{\infty} \mathbf{C}_\nu e^{-i\nu\omega}$ and $\mathbf{I}_{n,\epsilon}(\omega)$ is the periodogram based on $\underline{\epsilon}_1, \ldots, \underline{\epsilon}_n$. Now, we consider these three expressions seperately.

Theorem 2 in Hannan (1970, p.248) indicates that $\|\mathbf{I}_n(\omega) - \mathbf{C}(\omega)\mathbf{I}_{n,\epsilon}(\omega)\overline{\mathbf{C}(\omega)}^T\| = O_P(n^{-1/2})$ uniformly in ω and the supremum in (2.26) and in (2.28) vanish asymptotically in probability by standard arguments. Using again Taylor expansion for

$\mathbf{C}(\omega_k)$ the supremum in (2.27) can be bounded by

$$\sup_{\omega} \| \sum_{j_1,j_2=0}^{d} \mathbf{C}^{(j_1)}(\omega) \left(\frac{1}{nh^{d+1}} \sum_{k=-N}^{N} K^{(d)} \left(\frac{\omega-\omega_k}{h} \right) \frac{(\omega_k-\omega)^{s+j_1+j_2}}{j_1!j_2!} \mathbf{I}_{n,\epsilon}(\omega_k) \right)$$
$$\times \overline{\mathbf{C}^{(j_2)}(\omega)}^T - \frac{(-1)^d \mathbf{f}^{(d-s)}(\omega)}{(d-s)!} \frac{1}{nh} \sum_{k=-N}^{N} K^{(d)} \left(\frac{\omega-\omega_k}{h} \right) \left(\frac{\omega-\omega_k}{h} \right)^d \| + O_P(h).$$

Now, for instance, a multivariate version of Theorem 5.9.1 in Brillinger (1981) and following the approach of Franke and Härdle (1992) for proving Theorem A1 yield the claimed uniform convergence in probability of $\widetilde{\mathbf{Q}}(\omega)$ as n tends to infinity. Here, $(nh^6)^{-1} = O(1)$ has to be satisfied in comparison to Franke and Härdle, where no derivatives are estimated. □

2.7.2 Sample mean

Proof of Theorem 2.5.1

Since convergence in d_2-metric is equivalent to weak convergence and convergence of the first two moments [compare Bickel and Freedman (1981), Lemma 8.3], it suffices to show

$$Var^+(\sqrt{n}\ \underline{\overline{X}}^*) \to 2\pi\mathbf{f}(0),$$

where Var^+ is the conditional variance given $\underline{X}_1, \ldots, \underline{X}_n$ and

$$\mathcal{L}\{\sqrt{n}\ \underline{\overline{X}}^* | \underline{X}_1, \ldots, \underline{X}_n\} \Rightarrow \mathcal{N}(0, 2\pi\mathbf{f}(0))$$

in probability, respectively. Recall, it holds $Var(\sqrt{n}\underline{\overline{X}}) \to 2\pi\mathbf{f}(0)$ as $n \to \infty$ and $\sqrt{n}\underline{\overline{X}} \Rightarrow \mathcal{N}(0, 2\pi\mathbf{f}(0))$ [compare Brockwell and Davis (1991), p.406]. Straightforward calculation yields

$$\begin{aligned}
\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \underline{X}_t^* &= \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \sqrt{\frac{2\pi}{n}} \sum_{j=-[n/2]}^{[n/2]} \widetilde{\mathbf{Q}}(\omega_j) \underline{J}_n^+(\omega_j) e^{it\omega_j} \\
&= \frac{\sqrt{2\pi}}{n} \sum_{j=-[n/2]}^{[n/2]} \widetilde{\mathbf{Q}}(\omega_j) \underline{J}_n^+(\omega_j) \sum_{t=1}^{n} e^{it\omega_j} \\
&= \sqrt{2\pi}\widetilde{\mathbf{Q}}(0)\underline{J}_n^+(0) \\
&= \widetilde{\mathbf{Q}}(0) \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \underline{X}_t^+ \right)
\end{aligned}$$

and for the covariance matrix, we get immediately

$$Var^+ \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \underline{X}_t^* \right) = \widetilde{\mathbf{Q}}(0) Var^+ \left(\frac{1}{\sqrt{n}} \sum_{t=1}^{n} \underline{X}_t^+ \right) \widetilde{\mathbf{Q}}(0)^T.$$

For this reason, the claimed convergence in Mallows' metric follows from

$$\mathcal{L}\{\sqrt{n}\ \overline{\underline{X}}^{+}|\underline{X}_1,\ldots,\underline{X}_n\} \Rightarrow \mathcal{N}(0, 2\pi\mathbf{f}_{AR}(0)) \tag{2.29}$$

in probability, because, by construction, $2\pi\mathbf{Q}(0)\mathbf{f}_{AR}(0)\mathbf{Q}(0)^T = 2\pi\mathbf{f}(0)$. Using the Cramér-Wold device, assertion (2.29) results from an adequate CLT, e.g. for weakly dependent random variables as derived by Neumann and Paparoditis (2008, Theorem 6.1), which is well suited for the bootstrap. Thereby, we employ the convergence rate

$$\sup_{\nu\in\mathbb{N}_0}\|\widehat{\mathbf{C}}_\nu(p) - \mathbf{C}_\nu(p)\| = \frac{1}{r^\nu}O_P(n^{-1/2}), \tag{2.30}$$

for some $r > 1$ which was established by Kreiss (1984, p.7) for the coefficient matrices $\widehat{\mathbf{C}}_\nu(p)$, $\nu \in \mathbb{N}_0$ of the causal representation

$$\underline{X}_t^{+} = \sum_{\nu=0}^{\infty}\widehat{\mathbf{C}}_\nu(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\underline{\epsilon}_{t-\nu}^{+} \tag{2.31}$$

of the autoregressive fit of order p, using a multidimensional version of Cauchy's inequality for holomorphic functions [compare Kreiss and Franke (1992), Lemma 2.2 in the univariate case]. □

2.7.3 Spectral density

Proof of Theorem 2.5.2

To prove the Theorem, it is more convenient to use the *vec*-operator that creates a column vector by stacking the columns of a matrix below one another and to show the sufficient assertion

$$d_2\left\{\mathcal{L}(\sqrt{nb}vec\left(\left[\widehat{\mathbf{f}}(\omega_1) - \mathbf{f}(\omega_1)|\cdots|\widehat{\mathbf{f}}(\omega_s) - \mathbf{f}(\omega_s)\right]\right)),\right.$$
$$\left.\mathcal{L}(\sqrt{nb}vec\left(\left[\widehat{\mathbf{f}}^*(\omega_1) - \widetilde{\mathbf{f}}(\omega_1)|\cdots|\widehat{\mathbf{f}}^*(\omega_s) - \widetilde{\mathbf{f}}(\omega_s)\right]\right)|\underline{X}_1,\ldots,\underline{X}_n)\right\} \to 0$$

in probability. By Lemma 8.8 of Bickel and Freedman (1981), we can split the squared Mallows' metric in a variance part $V_n^2(\omega)$ and a squared bias part $b_n^2(\omega)$, where

$$\begin{aligned} & V_n^2(\omega) \\ = \ & d_2\left\{\mathcal{L}(\sqrt{nb}vec\left(\left[\widehat{\mathbf{f}}(\omega_1) - E[\widehat{\mathbf{f}}(\omega_1)]|\cdots|\widehat{\mathbf{f}}(\omega_s) - E[\widehat{\mathbf{f}}(\omega_s)]\right]\right)),\right. \\ & \left.\mathcal{L}(\sqrt{nb}vec\left(\left[\widehat{\mathbf{f}}^*(\omega_1) - E^+[\widehat{\mathbf{f}}^*(\omega_1)]|\cdots|\widehat{\mathbf{f}}^*(\omega_s) - E^+[\widehat{\mathbf{f}}^*(\omega_s)]\right]\right)|\underline{X}_1,\ldots,\underline{X}_n)\right\} \end{aligned}$$

and

$$\begin{aligned} b_n^2(\omega) = \ & nb\|vec\left(\left[E[\widehat{\mathbf{f}}(\omega_1)] - \mathbf{f}(\omega_1)|\cdots|E[\widehat{\mathbf{f}}(\omega_s)] - \mathbf{f}(\omega_s)\right]\right) \\ & -vec\left(\left[E^+[\widehat{\mathbf{f}}^*(\omega_1)] - \widetilde{\mathbf{f}}(\omega_1)|\cdots|E^+[\widehat{\mathbf{f}}^*(\omega_s)] - \widetilde{\mathbf{f}}(\omega_s)\right]\right)\|^2 \end{aligned}$$

and by Lemma 8.3 of the same paper, convergence in the d_2-metric is equivalent to weak convergence and convergence of the first two moments. The latter two follow from Lemma 2.7.2 and the weak convergence is a consequence of Lemma 2.7.3, so that $V_n^2(\omega) = o_P(1)$ holds. Recall that

$$nbCov(\widehat{f}_{jk}(\omega), \widehat{f}_{lm}(\lambda)) \\ \to \begin{cases} \{f_{jl}(\omega)f_{mk}(\omega) + f_{jm}(\omega)f_{lk}(\omega)\}\frac{1}{2\pi}\int K^2(u)du, & \omega = \lambda \in \{0, \pi\} \\ f_{jl}(\omega)f_{mk}(\omega)\frac{1}{2\pi}\int K^2(u)du, & 0 < \omega = \lambda < \pi \\ 0, & \omega \neq \lambda \end{cases} \tag{2.32}$$

and

$$\sqrt{nb}(\widehat{f}_{jk}(\omega_l) - E[\widehat{f}_{jk}(\omega_l)] : j, k = 1, \ldots, r; l = 1, \ldots, s)$$

is asymptotically (complex) normally distributed with mean vector $\underline{0}$ and covariance matrix obtained by (2.32) [compare Hannan (1970), Theorem 9, p. 280 and Theorem 11, p. 289 for a different, but asymptotically equivalent estimator]. Note that assumption (5.2) in Hannan (1970) is avoided in this context. Finally, the required convergence of $b_n^2(\omega)$ results from

$$\sqrt{nb}(E[\widehat{\mathbf{f}}(\omega)] - \mathbf{f}(\omega)) \to \frac{C}{4\pi}\mathbf{f}''(\omega)\int K(u)u^2du$$

for $nb^5 \to C^2 \geq 0$ as $n \to \infty$ and Lemma 2.7.4 below.

Lemma 2.7.2 (Covariance structure)**.**
Assume (A1), (A2), (K1) and (B2). For $j, k, l, m \in \{1, \ldots, r\}$ and $\omega, \lambda \in [0, \pi]$, the following convergence in probability holds true:

$$nbCov^+(\widehat{f}^*_{jk}(\omega), \widehat{f}^*_{lm}(\lambda)) \\ \to \begin{cases} \{f_{jl}(\omega)f_{mk}(\omega) + f_{jm}(\omega)f_{lk}(\omega)\}\frac{1}{2\pi}\int K^2(u)du, & \omega = \lambda \in \{0, \pi\} \\ f_{jl}(\omega)f_{mk}(\omega)\frac{1}{2\pi}\int K^2(u)du, & 0 < \omega = \lambda < \pi \\ 0, & \omega \neq \lambda \end{cases}$$

Proof.
We consider the case $\omega = \lambda \in [0, \pi]$ only. Let $k_1, k_2, h_1, h_2 \in \{1, \ldots, r\}$, then insertion

and straightforward calculation yields

$$
\begin{aligned}
& nbCov^{+}(\widehat{f}^{*}_{k_1h_1}(\omega), \widehat{f}^{*}_{k_2h_2}(\omega)) \\
= & \frac{b}{n} \sum_{j_1,j_2=-N}^{N} K_b(\omega-\omega_{j_1})K_b(\omega-\omega_{j_2}) \\
& \times \sum_{m_1,m_2,m_3,m_4=1}^{r} \widetilde{q}_{k_1m_1}(\omega_{j_1})\overline{\widetilde{q}_{h_1m_2}(\omega_{j_1})\widetilde{q}_{k_2m_3}(\omega_{j_2})}\widetilde{q}_{h_2m_4}(\omega_{j_2}) \\
& \times \left(E^{+}[I^{+}_{n,m_1m_2}(\omega_{j_1})\overline{I^{+}_{n,m_3m_4}(\omega_{j_2})}] - E^{+}[I^{+}_{n,m_1m_2}(\omega_{j_1})]E^{+}[\overline{I^{+}_{n,m_3m_4}(\omega_{j_2})}]\right) \\
= & \frac{b}{n} \sum_{j_1,j_2=-N}^{N} K_b(\omega-\omega_{j_1})K_b(\omega-\omega_{j_2}) \\
& \times \sum_{m_1,m_2,m_4,m_3=1}^{r} \widetilde{q}_{k_1m_1}(\omega_{j_1})\overline{\widetilde{q}_{h_1m_2}(\omega_{j_1})}\widetilde{q}_{h_2m_4}(\omega_{j_2})\overline{\widetilde{q}_{k_2m_3}(\omega_{j_2})}\frac{1}{4\pi^2n^2} \qquad (2.33) \\
& \times \sum_{s,t,v,u=1}^{n} \sum_{\nu_1,\nu_2,\nu_4,\nu_3=0}^{\infty} \sum_{\mu_1,\mu_2,\mu_4,\mu_3=1}^{r} \left(\widehat{\mathbf{C}}_{\nu_1}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_1\mu_1} \left(\widehat{\mathbf{C}}_{\nu_2}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_2\mu_2} \\
& \times \left(\widehat{\mathbf{C}}_{\nu_4}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_4\mu_4} \left(\widehat{\mathbf{C}}_{\nu_3}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_3\mu_3} e^{-i(s-t)\omega_{j_1}}e^{-i(v-u)\omega_{j_2}} \\
& \times \left(E^{+}[\epsilon^{+}_{s-\nu_1,\mu_1}\epsilon^{+}_{t-\nu_2,\mu_2}\epsilon^{+}_{v-\nu_4,\mu_4}\epsilon^{+}_{u-\nu_3,\mu_3}] - E^{+}[\epsilon^{+}_{s-\nu_1,\mu_1}\epsilon^{+}_{t-\nu_2,\mu_2}]E^{+}[\epsilon^{+}_{v-\nu_4,\mu_4}\epsilon^{+}_{u-\nu_3,\mu_3}]\right).
\end{aligned}
$$

Here, for the first equality we used $\mathbf{I}^{*}_{n}(\omega) = \widetilde{\mathbf{Q}}(\omega)\mathbf{I}^{+}_{n}(\omega)\overline{\widetilde{\mathbf{Q}}(\omega)}^{T}$ and the second results from inserting for the periodogram and (2.31). Because of the identity next to (2.41)-(2.43) we can deal with those three summands separately. Initially, we consider (2.43). Now, $\sum_{t=1}^{n} e^{it\omega} = 0$ if $\omega \neq 0$ and n otherwise causes the sum over j_2 in (2.33) to collapse and a rearrangement yields

$$
\begin{aligned}
& \frac{b}{n} \sum_{j=-N}^{N} K_b^2(\omega-\omega_j) \left(\sum_{m_1,m_3=1}^{r} \widetilde{q}_{k_1m_1}(\omega_j) \sum_{\mu_1=1}^{r} \frac{1}{2\pi} \sum_{\nu_1=0}^{\infty} \left(\widehat{\mathbf{C}}_{\nu_1}\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_1\mu_1} e^{-i\nu_1\omega_j} \right. \\
& \left. \times \sum_{\nu_3=0}^{\infty} \left(\widehat{\mathbf{C}}_{\nu_3}\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_3\mu_1} e^{i\nu_3\omega_j}\overline{\widetilde{q}_{k_2m_3}(\omega_j)} \right) \left(\sum_{m_4,m_2=1}^{r} \widetilde{q}_{h_2m_4}(\omega_j) \sum_{\mu_4=1}^{r} \frac{1}{2\pi} \right. \\
& \left. \times \sum_{\nu_4=0}^{\infty} \left(\widehat{\mathbf{C}}_{\nu_4}\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_4\mu_4} e^{-i\nu_4\omega_j} \sum_{\nu_2=0}^{\infty} \left(\widehat{\mathbf{C}}_{\nu_2}\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_2\mu_4} e^{i\nu_2\omega_j}\overline{\widetilde{q}_{h_1m_2}(\omega_j)} \right) \\
= & \frac{b}{n} \sum_{j=-N}^{N} K_b^2(\omega-\omega_j) \left(\widetilde{\mathbf{Q}}(\omega_j)\widehat{\mathbf{f}}_{\mathbf{AR}}(\omega_j)\overline{\widetilde{\mathbf{Q}}(\omega_j)}^{T}\right)_{k_1k_2} \left(\widetilde{\mathbf{Q}}(\omega_j)\widehat{\mathbf{f}}_{\mathbf{AR}}(\omega_j)\overline{\widetilde{\mathbf{Q}}(\omega_j)}^{T}\right)_{h_2h_1}.
\end{aligned}
$$

Because of the uniform convergence in ω of the quantities $\widetilde{\mathbf{Q}}(\omega)$ and $\widehat{\mathbf{f}}_{AR}(\omega)$, the last sum is asymptotically equal to

$$
\begin{aligned}
& \frac{b}{n}\sum_{j=-N}^{N} K_b^2(\omega-\omega_j)\left(\mathbf{Q}(\omega_j)\mathbf{f}_{\mathbf{AR}}(\omega_j)\overline{\mathbf{Q}(\omega_j)}^T\right)_{k_1k_2}\left(\mathbf{Q}(\omega_j)\mathbf{f}_{\mathbf{AR}}(\omega_j)\overline{\mathbf{Q}(\omega_j)}^T\right)_{h_2h_1} \\
= & \frac{1}{nb}\sum_{j=-N}^{N} K^2\left(\frac{\omega-\omega_j}{b}\right) f_{k_1k_2}(\omega_j)f_{h_2h_1}(\omega_j),
\end{aligned}
$$

where we used the correcting property of $\widetilde{\mathbf{Q}}(\omega)$. Concerning assumption $(A1)$, the spectral density $\mathbf{f}(\omega)$ is componentwise differentiable with bounded derivative. For this reason, Taylor expansions of $f_{k_1k_2}(\omega_j)$ and $f_{h_2h_1}(\omega_j)$ plus the converging Riemann sum yield

$$
\begin{aligned}
& \frac{1}{nb}\sum_{j=-N}^{N} K^2\left(\frac{\omega-\omega_j}{b}\right) f_{k_1k_2}(\omega_j)f_{h_2h_1}(\omega_j) + o_P(1) \\
= \quad & f_{k_1k_2}(\omega)f_{h_2h_1}(\omega)\frac{1}{2\pi}\left(\frac{2\pi}{nb}\sum_{j=-N}^{N} K^2\left(\frac{\omega-\omega_j}{b}\right)\right) + O_P(b) + o_P(1) \\
\to \quad & f_{k_1k_2}(\omega)f_{h_2h_1}(\omega)\frac{1}{2\pi}\int_{-\pi}^{\pi} K^2(x)dx
\end{aligned}
$$

in probability. Arguments are similar for the term related to (2.42) and we get

$$
f_{k_1h_2}(\omega)f_{k_2h_1}(\omega)\frac{1}{2\pi}\left(\frac{2\pi}{nb}\sum_{j=-N}^{N} K\left(\frac{\omega-\omega_j}{b}\right)K\left(\frac{\omega+\omega_j}{b}\right)\right) + O_P(b) + o_P(1),
$$

where the involved Riemann sum converges to zero for $\omega \in (0,\pi)$ and to $\frac{1}{2\pi}\int K^2(u)du$ for $\omega \in \{0,\pi\}$ as required. It remains to check (2.41) concerning its asymptotic behaviour. Inserting (2.41) in equation (2.33) and standard calculations result in an $o_P(1)$ term that, for this reason, does not play a role asymptotically. This completes the proof. $\square$

Lemma 2.7.3 (Asymptotic normality)**.**
Assume (A1), (A2), (K1), (K2) and (B2). Then, the following assertion holds true:

$$
\begin{aligned}
& \mathcal{L}\left(\sqrt{nb}\,\mathrm{vec}\left(\left[\widehat{\mathbf{f}}^*(\omega_1) - E^+[\widehat{\mathbf{f}}^*(\omega_1)]|\cdots|\widehat{\mathbf{f}}^*(\omega_s) - E^+[\widehat{\mathbf{f}}^*(\omega_s)]\right]\right)\Big|\, \underline{X}_1,\ldots,\underline{X}_n\right) \\
\Rightarrow \quad & \mathcal{N}^{\mathbb{C}}(\underline{0},\mathbf{W})
\end{aligned}
$$

in probability, where $\mathcal{N}^{\mathbb{C}}$ *denotes a complex normal distribution [cf. Brillinger (1981), p.89] and the asymptotic covariance matrix* $\mathbf{W}$ *is obtained by the results of Lemma 2.7.2.*

Proof.
Let $\underline{c} = (\underline{c}^{(1)T}, \ldots, \underline{c}^{(s)T})^T \in \mathbb{C}^{sr^2}$ with $\underline{c}^{(l)} \in \mathbb{C}^{r^2}$, $l = 1, \ldots, s$. Using the Cramér-Wold device, applied to complex-valued random variables, it suffices to show asymptotic normality for

$$\begin{aligned} & \underline{c}^T \sqrt{nb}\, vec\left(\left[\widehat{\mathbf{f}}^*(\omega_1) - E^+[\widehat{\mathbf{f}}^*(\omega_1)] | \cdots | \widehat{\mathbf{f}}^*(\omega_s) - E^+[\widehat{\mathbf{f}}^*(\omega_s)]\right]\right) \\ = & \sum_{l=1}^{s} \underline{c}^{(l)\ T} \sqrt{nb}\, vec\left(\left[\widehat{\mathbf{f}}^*(\omega_l) - E^+[\widehat{\mathbf{f}}^*(\omega_l)]\right]\right). \end{aligned}$$

For this reason, without loss of generality, we can restrict our considerations to the case $s = 1$. Analogue to Theorem 2 in Hannan (1970, p.248), it holds

$$\mathbf{I}_n^+(\omega) = \left(\sum_{\nu=0}^{\infty} \widehat{\mathbf{C}}_\nu(p) e^{-i\nu\omega}\right) \widehat{\mathbf{\Sigma}}^{1/2}(p) \mathbf{I}_{n,\epsilon^+}(\omega) \widehat{\mathbf{\Sigma}}^{1/2}(p)^T \overline{\left(\sum_{\nu=0}^{\infty} \widehat{\mathbf{C}}_\nu(p) e^{-i\nu\omega}\right)}^T$$

up to an $O_{P^*}(\frac{1}{\sqrt{n}})$ term, where $\widehat{\mathbf{\Sigma}}^{1/2}(p)$ is defined in Step 1 in Section 2.4 and $\mathbf{I}_{n,\epsilon^+}(\omega)$ is the periodogram based on the bootstrap residuals $\underline{\epsilon}_1^+, \ldots, \underline{\epsilon}_n^+$. Using this formula and

$$\mathbf{I}_n^*(\omega) = \widetilde{\mathbf{Q}}(\omega) \mathbf{I}_n^+(\omega) \overline{\widetilde{\mathbf{Q}}(\omega)}^T,$$

we get

$$\begin{aligned} & \sqrt{nb}(\widehat{\mathbf{f}}^*(\omega) - E^+[\widehat{\mathbf{f}}^*(\omega)]) \\ = & \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j) \widehat{\mathbf{M}}(\omega_j) \left(\mathbf{I}_{n,\epsilon^+}(\omega_j) - \frac{1}{2\pi}\mathbf{I}_r\right) \overline{\widehat{\mathbf{M}}(\omega_j)}^T + o_{P^*}(1), \end{aligned}$$

where $\widehat{\mathbf{M}}(\omega) = \widetilde{\mathbf{Q}}(\omega)(\sum_{\nu=0}^{\infty} \widehat{\mathbf{C}}_\nu(p) e^{-i\nu\omega}) \widehat{\mathbf{\Sigma}}^{1/2}(p)$. Thanks to a multivariate analogue to Theorem 5.9.1 in Brillinger (1981), instead of the first term on the right-hand side of the above equality, we may consider the asymptotically equivalent statistic

$$\begin{aligned} & \sqrt{nb}\frac{1}{2\pi} \int_{-\pi}^{\pi} K_b(\omega - x) \mathbf{M}(x) \left(\mathbf{I}_{n,\epsilon^+}(x) - \frac{1}{2\pi}\mathbf{I}_r\right) \overline{\mathbf{M}(x)}^T dx \\ = & \sqrt{nb}\frac{1}{2\pi} \int_{-\pi}^{\pi} K(u) \mathbf{M}(\omega - ub) \left(\mathbf{I}_{n,\epsilon^+}(\omega - ub) - \frac{1}{2\pi}\mathbf{I}_r\right) \overline{\mathbf{M}(\omega - ub)}^T du \\ = & \mathbf{M}(\omega) \left(\sqrt{nb}\frac{1}{2\pi} \int_{-\pi}^{\pi} K(u) \left(\mathbf{I}_{n,\epsilon^+}(\omega - ub) - \frac{1}{2\pi}\mathbf{I}_r\right) du\right) \overline{\mathbf{M}(\omega)}^T \qquad (2.34) \\ & + \mathbf{D}_{n,1}^+(\omega) + \mathbf{D}_{n,2}^+(\omega), \end{aligned}$$

where $\mathbf{M}(\omega) = \mathbf{Q}(\omega) \left(\sum_{\nu=0}^{\infty} \mathbf{C}_\nu(p) e^{-i\nu\omega}\right) \mathbf{\Sigma}^{1/2}(p)$ is the limit in probability of $\widehat{\mathbf{M}}(\omega)$

and the quantities $\mathbf{D}_{n,1}^+(\omega)$ and $\mathbf{D}_{n,2}^+(\omega)$ are defined as follows:

$$\begin{aligned}\mathbf{D}_{n,1}^+(\omega) &= \sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi}K(u)\\ &\quad\times(\mathbf{M}(\omega-ub)-\mathbf{M}(\omega))\left(\mathbf{I}_{n,\epsilon^+}(\omega-ub)-\frac{1}{2\pi}\mathbf{I}_r\right)\overline{\mathbf{M}(\omega-ub)}^T du,\\ \mathbf{D}_{n,2}^+(\omega) &= \sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi}K(u)\\ &\quad\times\mathbf{M}(\omega-ub)\left(\mathbf{I}_{n,\epsilon^+}(\omega-ub)-\frac{1}{2\pi}\mathbf{I}_r\right)\overline{(\mathbf{M}(\omega-ub)-\mathbf{M}(\omega))}^T du.\end{aligned}$$

For the components $D_{n,k}^+(i,j)(\omega)$ of $\mathbf{D}_{n,k}^+(\omega)$, $k\in\{1,2\}$, straightforward calculations yield $E^+[D_{n,k}^+(i,j)(\omega)]=0$ and

$$E^+[|D_{n,k}^+(i,j)(\omega)|^2]=O_P(\max_{i,j=1,\ldots,r}\{|M_{i,j}(\omega-ub)-M_{i,j}(\omega)|^2\})=O_P(b^2)$$

for all $i,j\in\{1,\ldots,r\}$, where the last equality follows from the Lipschitz-continuity of $\mathbf{M}$, which is a consequence of this property fulfilled by $\mathbf{Q}(\omega)$ and $\sum_{\nu=0}^{\infty}\mathbf{C}_\nu(p)e^{-i\nu\omega}$. Due to the formula $vec(\mathbf{ABC})=(\mathbf{C}^T\otimes\mathbf{A})vec(\mathbf{B})$ for appropriate matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ [cf. Lütkepohl (2005), p.662], the term in (2.34) becomes

$$(\mathbf{M}(\omega)\otimes\overline{\mathbf{M}(\omega)})vec\left(\sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi}K(u)\left(\mathbf{I}_{n,\epsilon^+}(\omega-ub)-\frac{1}{2\pi}\mathbf{I}_r\right)du\right)$$

and it remains to show asymptotic normality for the part in big outer parentheses above. Plugging-in the expression $\mathbf{I}_{n,\epsilon^+}(\omega)=\frac{1}{2\pi}\sum_{s=-n+1}^{n-1}\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)e^{-is\omega}$, where

$$\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)=\begin{cases}\frac{1}{n}\sum_{t=1}^{n-s}\underline{\epsilon}_{t+s}^+\underline{\epsilon}_t^{+T}, & s\geq 0\\ \frac{1}{n}\sum_{t=1-s}^{n}\underline{\epsilon}_{t+s}^+\underline{\epsilon}_t^{+T}, & s<0\end{cases}, \tag{2.35}$$

we get

$$\begin{aligned}&\sqrt{nb}\frac{1}{2\pi}\int_{-\pi}^{\pi}K(u)\left(\mathbf{I}_{n,\epsilon^+}(\omega-ub)-\frac{1}{2\pi}\mathbf{I}_r\right)du\\ =&\sqrt{nb}\int_{-\pi}^{\pi}K(u)\frac{1}{4\pi^2}\sum_{s=1}^{n-1}\left(\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)e^{-is(\omega-ub)}+\widehat{\mathbf{\Gamma}}_{\epsilon^+}(-s)e^{is(\omega-ub)}\right)du\\ &+\sqrt{nb}\frac{1}{4\pi^2}(\widehat{\mathbf{\Gamma}}_{\epsilon^+}(0)-\mathbf{I}_r)\int_{-\pi}^{\pi}K(u)du,\end{aligned}$$

where the second term is $O_{P^*}(\sqrt{b})=o_{P^*}(1)$. Using the Fourier transform k of K and its symmetry, the first term can be written as

$$\frac{1}{4\pi^2}\sqrt{nb}\sum_{s=1}^{n-1}k(sb)\left(\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)e^{-is\omega}+\widehat{\mathbf{\Gamma}}_{\epsilon^+}(-s)e^{is\omega}\right).$$

Ignoring the factor $\frac{1}{4\pi^2}$, we can split this expression to obtain

$$\sqrt{nb}\sum_{s=1}^{c_n} k(sb)\left(\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(s)e^{-is\omega}+\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(-s)e^{is\omega}\right) \tag{2.36}$$

$$+\sqrt{nb}\sum_{s=1}^{c_n} k(sb)\left(\left(\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)-\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(s)\right)e^{-is\omega}+\left(\widehat{\mathbf{\Gamma}}_{\epsilon^+}(-s)-\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(-s)\right)e^{is\omega}\right) \tag{2.37}$$

$$+\sqrt{nb}\sum_{s=c_n+1}^{n-1} k(sb)\left(\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)e^{-is\omega}+\widehat{\mathbf{\Gamma}}_{\epsilon^+}(-s)e^{is\omega}\right), \tag{2.38}$$

where $(c_n, n\in\mathbb{N})\subset\mathbb{N}$ satisfies $c_n=o(n)$ as well as $c_n\to\infty$ as $n\to\infty$ and the summation in $\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(s)$ is from 1 to n compared to the definition of $\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)$ in (2.35). Next we show that (2.37) and (2.38) vanish asymptotically. We prove this only for the parts with positive lags s. For the (h,j)-th component of (2.37), we get

$$\begin{aligned}
&E^+[|\sqrt{nb}\sum_{s=1}^{c_n} k(sb)\left(\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)-\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(s)\right)_{h,j}e^{-is\omega}|^2]\\
=\;& nbE^+[|\sum_{s=1}^{c_n} k(sb)\frac{1}{n}\sum_{t=n-s+1}^{n}\epsilon^+_{t+s,h}\epsilon^+_{t,j}e^{-is\omega}|^2]\\
\le\;& nb\sum_{s=1}^{c_n} k^2(sb)\frac{s}{n^2}.
\end{aligned}$$

The last term is bounded by $\frac{c_n b}{n}\sum_{s=1}^{c_n} k^2(sb)$, which, defining $m_n=[\frac{1}{b}]$, is asymptotically equivalent to

$$\frac{c_n}{n}\frac{1}{m_n}\sum_{s=1}^{c_n} k^2\left(\frac{s}{m_n}\right) \;\cong\; \frac{c_n}{n}\int_0^{c_n} k^2(x)dx\to 0,$$

where $\int k(u)^2du<\infty$ and $c_n=o(n)$ are used. Similarly, for the (h,j)-th component of (2.38), we get

$$\begin{aligned}
E^+[|\sqrt{nb}\sum_{s=c_n+1}^{n-1} k(sb)\widehat{\mathbf{\Gamma}}_{\epsilon^+}(s)_{h,j}e^{-is\omega}|^2] \;&=\; nbE^+[|\sum_{s=c_n+1}^{n-1} k(sb)\frac{1}{n}\sum_{t=1}^{n-s}\epsilon^+_{t+s,h}\epsilon^+_{t,j}e^{-is\omega}|^2]\\
&\le\; b\sum_{s=c_n+1}^{n-1} k^2(sb)
\end{aligned}$$

and the last sum is asymptotically equivalent to

$$\frac{1}{m_n}\sum_{s=c_n+1}^{n-1} k^2\left(\frac{s}{m_n}\right) \;\cong\; \int_{c_n}^{\infty} k^2(x)dx\to 0.$$

Using expression (2.36), now, we define the quantity $\mathbf{W}_{t,n}^{+}$ by

$$\begin{aligned}&\sqrt{nb}\sum_{s=1}^{c_n} k(sb)\left(\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(s)e^{-is\omega}+\widetilde{\mathbf{\Gamma}}_{\epsilon^+}(-s)e^{is\omega}\right)\\ =&\sum_{t=1}^{n}\sqrt{\frac{b}{n}}\sum_{s=1}^{c_n}k(sb)\left(\underline{\epsilon}_{t+s}^{+}\underline{\epsilon}_{t}^{+T}e^{-is\omega}+\underline{\epsilon}_{t-s}^{+}\underline{\epsilon}_{t}^{+T}e^{is\omega}\right)=:\sum_{t=1}^{n}\mathbf{W}_{t,n}^{+}\end{aligned}$$

and, by the Cramér-Wold device, finally, it remains to show asymptotic (complex) normality of $\sum_{t=1}^{n}\underline{c}^{T}vec\left(\mathbf{W}_{t,n}^{+}\right)$ for all $\underline{c}\in\mathbb{C}^{r^2}$, which per definition of the complex normal distribution is equivalent to asymptotic (real) normality of

$$\sum_{t=1}^{n}\underline{c}^{T}vec\left([Re(\mathbf{W}_{t,n}^{+})|Im(\mathbf{W}_{t,n}^{+})]\right)=\sum_{t=1}^{n}\underline{c}^{(1)T}vec\left(Re(\mathbf{W}_{t,n}^{+})\right)+\underline{c}^{(2)T}vec\left(Im(\mathbf{W}_{t,n}^{+})\right)$$

for all $\underline{c}=(\underline{c}^{(1)T},\underline{c}^{(2)T})^{T}\in\mathbb{R}^{2r^2}$, where $Re(x)$ and $Im(x)$ denote the real and the imaginary part of a complex quantity x. These one-dimensional quantities can be treated standardly with Theorem 4 in Rosenblatt (1985, p.63) as done in Kreiss and Paparoditis (2003) for the univariate case to obtain asymptotic normality using the AAPB, which completes this proof. □

Lemma 2.7.4 (Bias term)**.**
Assume (A1), (A2), (A3), (K1), (K3) and (B3). If $nb^5\to C^2$ *with a constant* $C\geq 0$*, we get*

$$\sqrt{nb}(E^{+}[\widehat{\mathbf{f}}^{*}(\omega)]-\widetilde{\mathbf{f}}(\omega))\to\frac{C}{4\pi}\mathbf{f}''(\omega)\int K(u)u^2du$$

in probability, where $\mathbf{f}''(\omega)$ *is the (entrywise) second derivative in* ω *of the spectral density matrix* $\mathbf{f}$*.*

Proof.
Thanks to $|\frac{1}{n}\sum_{j=-N}^{N}K_b(\omega-\omega_j)-1|=O(\frac{1}{nb})$ uniformly in ω and $E^{+}[\mathbf{I}_n^{+}(\omega_j)]=\widehat{\mathbf{f}}_{AR}(\omega_j)+O(\frac{1}{n})$ uniformly in ω_j, at first, we get

$$\begin{aligned}&\sqrt{nb}(E^{+}[\widehat{\mathbf{f}}^{*}(\omega)]-\widetilde{\mathbf{f}}(\omega))\\ =&\sqrt{nb}\left(\frac{1}{n}\sum_{j=-N}^{N}K_b(\omega-\omega_j)\left(\widetilde{\mathbf{Q}}(\omega_j)\widehat{\mathbf{f}}_{AR}(\omega_j)\overline{\widetilde{\mathbf{Q}}(\omega_j)}^{T}-\widetilde{\mathbf{Q}}(\omega)\widehat{\mathbf{f}}_{AR}(\omega)\overline{\widetilde{\mathbf{Q}}(\omega)}^{T}\right)\right)\\ &+O_P\left(\frac{1}{\sqrt{nb}}\right)+O_P\left(\sqrt{b}\right).\end{aligned}$$

Now, the expression in inner round parentheses can be displayed in the following way:

$$\begin{aligned}
&\widetilde{\mathbf{Q}}(\omega_j)\widehat{\mathbf{f}}_{AR}(\omega_j)\overline{\widetilde{\mathbf{Q}}(\omega_j)}^T - \widetilde{\mathbf{Q}}(\omega)\widehat{\mathbf{f}}_{AR}(\omega)\overline{\widetilde{\mathbf{Q}}(\omega)}^T \\
= \;& (\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))\widehat{\mathbf{f}}_{AR}(\omega)\overline{\widetilde{\mathbf{Q}}(\omega)}^T + \widetilde{\mathbf{Q}}(\omega)(\widehat{\mathbf{f}}_{AR}(\omega_j) - \widehat{\mathbf{f}}_{AR}(\omega))\overline{\widetilde{\mathbf{Q}}(\omega)}^T \\
&+\widetilde{\mathbf{Q}}(\omega)\widehat{\mathbf{f}}_{AR}(\omega)\overline{(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))}^T + (\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))(\widehat{\mathbf{f}}_{AR}(\omega_j) - \widehat{\mathbf{f}}_{AR}(\omega))\overline{\widetilde{\mathbf{Q}}(\omega)}^T \\
&+(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))\widehat{\mathbf{f}}_{AR}(\omega)\overline{(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))}^T \\
&+\widetilde{\mathbf{Q}}(\omega)(\widehat{\mathbf{f}}_{AR}(\omega_j) - \widehat{\mathbf{f}}_{AR}(\omega))\overline{(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))}^T \\
&+(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))(\widehat{\mathbf{f}}_{AR}(\omega_j) - \widehat{\mathbf{f}}_{AR}(\omega))\overline{(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))}^T \\
= \;& \widehat{\mathbf{D}}_{1,j} + \widehat{\mathbf{D}}_{2,j} + \widehat{\mathbf{D}}_{3,j} + \widehat{\mathbf{D}}_{4,j} + \widehat{\mathbf{D}}_{5,j} + \widehat{\mathbf{D}}_{6,j} + \widehat{\mathbf{D}}_{7,j},
\end{aligned}$$

with an obvious notation for $\widehat{\mathbf{D}}_{k,j}$, $k = 1, \ldots, 7$. Note, because of the chain rule, for the second (componentwise) derivative of $\mathbf{f}(\omega)$, it holds

$$\begin{aligned}
\mathbf{f}''(\omega) &= (\mathbf{Q}(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T)'' \\
&= \mathbf{Q}''(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T + \mathbf{Q}(\omega)\mathbf{f}''_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T + \mathbf{Q}(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}''(\omega)}^T \\
&\quad +2\mathbf{Q}'(\omega)\mathbf{f}'_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T + 2\mathbf{Q}'(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}'(\omega)}^T + 2\mathbf{Q}(\omega)\mathbf{f}'_{AR}(\omega)\overline{\mathbf{Q}'(\omega)}^T \\
&= \mathbf{D}_1 + \mathbf{D}_2 + \mathbf{D}_3 + \mathbf{D}_4 + \mathbf{D}_5 + \mathbf{D}_6
\end{aligned}$$

and the claimed convergence of $E^+[\widehat{\mathbf{f}}^*(\omega)] - \widetilde{\mathbf{f}}(\omega)$ follows from

$$\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)\widehat{\mathbf{D}}_{k,j} \to \frac{C}{4\pi}\mathbf{D}_k \int K(u)u^2 du, \quad k = 1, \ldots, 7 \tag{2.39}$$

in probability, where $\mathbf{D}_7$ is set equal to zero. Consider first $\widehat{\mathbf{D}}_{1,j}$. A Taylor expansion of $\widetilde{\mathbf{Q}}(\omega)$ delivers

$$\begin{aligned}
&\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega)) \\
= \;& \left(\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)\right)\widetilde{\mathbf{Q}}'(\omega) + \left(\frac{1}{2}\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)^2\right) \\
&\times\widetilde{\mathbf{Q}}''(\omega) + \left(\frac{1}{6}\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)^3\widetilde{\mathbf{Q}}'''(\widetilde{\omega}_j)\right)
\end{aligned}$$

with $\widetilde{\omega}_j$ between ω and ω_j. Due to $\int K(u)udu = 0$ we get $\frac{1}{nb}\sum_{j=-N}^{N} K(\frac{\omega-\omega_j}{b})(\frac{\omega_j-\omega}{b}) = O(\frac{1}{nb})$ and together with $nb^5 = O(1)$ the first summand vanishes. The third is $O_P(b)$

because of $\widetilde{\mathbf{Q}}'''(\omega) = O_P(1)$ uniformly in ω and disappears also. From $nb^5 \to C^2$ and Lemma 2.7.1, for the second term, we get

$$\left(\frac{\sqrt{nb^5}}{4\pi} \frac{2\pi}{nb} \sum_{j=-N}^{N} K\left(\frac{\omega - \omega_j}{b}\right) \left(\frac{\omega_j - \omega}{b}\right)^2 \right) \widetilde{\mathbf{Q}}''(\omega) \to \frac{C}{4\pi} \mathbf{Q}''(\omega) \int K(u) u^2 du,$$

which yields (2.39) for $k = 1$. The cases $k = 2$ and $k = 3$ can be treated analogously, where a Taylor expansion of $\widehat{\mathbf{f}}_{AR}(\omega)$ has to be used for $k = 2$. Now, consider $k \in \{4, 5, 6\}$. We prove only the case $k = 4$. Similar to calculations above, Taylor expansions of $\widetilde{\mathbf{Q}}(\omega)$ and $\widehat{\mathbf{f}}_{AR}(\omega)$, respectively, provide

$$\begin{aligned}
& \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))(\widehat{\mathbf{f}}_{AR}(\omega_j) - \widehat{\mathbf{f}}_{AR}(\omega)) \\
= & \left(\sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)^2 \right) \widetilde{\mathbf{Q}}'(\omega) \widehat{\mathbf{f}}'_{AR}(\omega) \\
& + \left(\frac{1}{2} \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)^3 \widetilde{\mathbf{Q}}''(\widetilde{\omega}) \right) \widehat{\mathbf{f}}'_{AR}(\omega) \\
& + \widetilde{\mathbf{Q}}'(\omega) \left(\frac{1}{2} \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)^3 \widehat{\mathbf{f}}''_{AR}(\widetilde{\omega}) \right) \\
& + \frac{1}{4} \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)^4 \widetilde{\mathbf{Q}}''(\widetilde{\omega}) \widehat{\mathbf{f}}''_{AR}(\widetilde{\omega}) \\
\to & \frac{C}{4\pi} \mathbf{Q}'(\omega) \mathbf{f}'_{AR}(\omega) \int K(u) u^2 du.
\end{aligned}$$

Finally, three times Taylor again yields

$$\begin{aligned}
& \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))(\widehat{\mathbf{f}}_{AR}(\omega_j) - \widehat{\mathbf{f}}_{AR}(\omega))\overline{(\widetilde{\mathbf{Q}}(\omega_j) - \widetilde{\mathbf{Q}}(\omega))}^T \\
= & \sqrt{\frac{b}{n}} \sum_{j=-N}^{N} K_b(\omega - \omega_j)(\omega_j - \omega)^3 \widetilde{\mathbf{Q}}'(\widetilde{\omega}) \widehat{\mathbf{f}}'_{AR}(\widetilde{\omega}) \overline{\widetilde{\mathbf{Q}}'(\widetilde{\omega})}^T
\end{aligned}$$

and the last sum vanishes asymptotically, because of $\int K(u) u^3 du = 0$. □

This concludes the proof of Theorem 2.5.2. □

2.7.4 Autocovariances

Proof of Theorem 2.5.3

Extending $\underline{X}_1^*, \ldots, \underline{X}_n^*$ cyclically to obtain $(\underline{X}_t^*, t \in \mathbb{Z})$, we can define

$$\widetilde{\boldsymbol{\Gamma}}^*(h) = \frac{1}{n}\sum_{t=1}^{n} \underline{X}_{t+h}^* \underline{X}_t^{*T}, \quad h \in \mathbb{Z}$$

and because of $E[\|\widehat{\boldsymbol{\Gamma}}^*(h) - \widetilde{\boldsymbol{\Gamma}}^*(h)\|] = O_P(\frac{1}{n})$ it suffices to show the assertion for the components of $\widetilde{\boldsymbol{\Gamma}}^*(h)$. Let $h_1, h_2, k_1, k_2 \in \{1, \ldots, r\}$ as well as $h, k \in \mathbb{Z}$, then insertion yields

$$\begin{aligned}
& nCov^+(\widetilde{\gamma}^*_{h_1h_2}(h) - E^+[\widetilde{\gamma}^*_{h_1h_2}(h)], \widetilde{\gamma}^*_{k_1k_2}(k) - E^+[\widetilde{\gamma}^*_{k_1k_2}(k)]) \\
= & \frac{4\pi^2}{n^3} \sum_{j_1,j_2,j_3,j_4=-N}^{N} \sum_{m_1,m_2,m_3,m_4=1}^{r} \widetilde{q}_{h_1m_1}(\omega_{j_1})\widetilde{q}_{h_2m_2}(\omega_{j_2})\widetilde{q}_{k_1m_3}(\omega_{j_3})\widetilde{q}_{k_2m_4}(\omega_{j_4}) \\
& \times \big(E^+[J^+_{n,m_1}(\omega_{j_1})J^+_{n,m_2}(\omega_{j_2})J^+_{n,m_3}(\omega_{j_3})J^+_{n,m_4}(\omega_{j_4})] \\
& \quad -E^+[J^+_{n,m_1}(\omega_{j_1})J^+_{n,m_2}(\omega_{j_2})]E^+[J^+_{n,m_3}(\omega_{j_3})J^+_{n,m_4}(\omega_{j_4})]\big) \\
& \times e^{ih\omega_{j_1}} e^{ik\omega_{j_3}} \sum_{s=1}^{n} e^{is(\omega_{j_1}+\omega_{j_2})} \sum_{t=1}^{n} e^{it(\omega_{j_3}+\omega_{j_4})}
\end{aligned}$$

and with straightforward calculation the last expression becomes

$$\begin{aligned}
& \frac{4\pi^2}{n} \sum_{j_1,j_2=-N}^{N} \sum_{m_1,m_2,m_3,m_4=1}^{r} \widetilde{q}_{h_1m_1}(\omega_{j_1})\overline{\widetilde{q}_{h_2m_2}(\omega_{j_1})}\widetilde{q}_{k_1m_3}(\omega_{j_2})\overline{\widetilde{q}_{k_2m_4}(\omega_{j_2})} \\
& \times \big(E^+[I^+_{n,m_1m_2}(\omega_{j_1})I^+_{n,m_3m_4}(\omega_{j_2})] - E^+[I^+_{n,m_1m_2}(\omega_{j_1})]E^+[I^+_{n,m_3m_4}(\omega_{j_2})]\big) \\
& \times e^{ih\omega_{j_1}} e^{ik\omega_{j_2}} \\
= & \frac{4\pi^2}{n} \sum_{j_1,j_2=-N}^{N} \sum_{m_1,m_2,m_3,m_4=1}^{r} \widetilde{q}_{h_1m_1}(\omega_{j_1})\overline{\widetilde{q}_{h_2m_2}(\omega_{j_1})}\widetilde{q}_{k_1m_3}(\omega_{j_2})\overline{\widetilde{q}_{k_2m_4}(\omega_{j_2})} \\
& \times \frac{1}{4\pi^2 n^2} \sum_{s,t,u,v=1}^{n} \sum_{\nu_1,\nu_2,\nu_3,\nu_4=0}^{\infty} \sum_{\mu_1,\mu_2,\mu_3,\mu_4=1}^{r} \Big(\widehat{\mathbf{C}}_{\nu_1}(p)\widehat{\boldsymbol{\Sigma}}^{1/2}(p)\Big)_{m_1\mu_1} \\
& \times \Big(\widehat{\mathbf{C}}_{\nu_2}(p)\widehat{\boldsymbol{\Sigma}}^{1/2}(p)\Big)_{m_2\mu_2} \Big(\widehat{\mathbf{C}}_{\nu_3}(p)\widehat{\boldsymbol{\Sigma}}^{1/2}(p)\Big)_{m_3\mu_3} \Big(\widehat{\mathbf{C}}_{\nu_4}(p)\widehat{\boldsymbol{\Sigma}}^{1/2}(p)\Big)_{m_4\mu_4} \\
& \times \big(E^+[\epsilon^+_{s-\nu_1,\mu_1}\epsilon^+_{t-\nu_2,\mu_2}\epsilon^+_{u-\nu_3,\mu_3}\epsilon^+_{v-\nu_4,\mu_4}] \\
& \quad\quad -E^+[\epsilon^+_{s-\nu_1,\mu_1}\epsilon^+_{t-\nu_2,\mu_2}]E^+[\epsilon^+_{u-\nu_3,\mu_3}\epsilon^+_{v-\nu_4,\mu_4}]\big) \\
& \times e^{-i(s-t)\omega_{j_1}} e^{-i(u-v)\omega_{j_2}} e^{ih\omega_{j_1}} e^{ik\omega_{j_2}}.
\end{aligned} \tag{2.40}$$

For the second equality from above we used $\sum_{t=1}^{n} e^{it\omega} = 0$ if $\omega \neq 0$ and n otherwise as well as the hermitian symmetry of $\underline{J}_n(\omega)$ and $\widetilde{\mathbf{Q}}(\omega)$. Inserting for the periodogram provides the third equation. Up to an eventually negligible term, it holds

$$
\begin{aligned}
&\sum_{s,t,u,v=1}^{n}\left(E[\epsilon^+_{s-\nu_1,\mu_1}\epsilon^+_{t-\nu_2,\mu_2}\epsilon^+_{u-\nu_3,\mu_3}\epsilon^+_{v-\nu_4,\mu_4}]-E[\epsilon^+_{s-\nu_1,\mu_1}\epsilon^+_{t-\nu_2,\mu_2}]E[\epsilon^+_{u-\nu_3,\mu_3}\epsilon^+_{v-\nu_4,\mu_4}]\right)\\
&e^{-i(s-t)\omega_{j_1}}e^{-i(u-v)\omega_{j_2}}\\
=\;& n\widetilde{\kappa}_4(p;\mu_1,\mu_2,\mu_3,\mu_4)e^{-i\nu_1\omega_{j_1}}e^{i\nu_2\omega_{j_1}}e^{-i\nu_3\omega_{j_2}}e^{i\nu_4\omega_{j_2}} && (2.41)\\
&+1(\mu_1=\mu_3)1(\mu_2=\mu_4)\left|\sum_{s=1}^{n}e^{-is(\omega_{j_1}+\omega_{j_2})}\right|^2 e^{i\nu_1\omega_{j_2}}e^{-i\nu_2\omega_{j_2}}e^{-i\nu_3\omega_{j_2}}e^{i\nu_4\omega_{j_2}} && (2.42)\\
&+1(\mu_1=\mu_4)1(\mu_2=\mu_3)\left|\sum_{s=1}^{n}e^{-is(\omega_{j_1}-\omega_{j_2})}\right|^2 e^{-i\nu_1\omega_{j_2}}e^{i\nu_2\omega_{j_2}}e^{-i\nu_3\omega_{j_2}}e^{i\nu_4\omega_{j_2}}, && (2.43)
\end{aligned}
$$

where $\widetilde{\kappa}_4(p;\mu_1,\mu_2,\mu_3,\mu_4)$ denotes the fourth-order cumulant between the *standardized* residuals $\epsilon^+_{t,\mu_1},\epsilon^+_{t,\mu_2},\epsilon^+_{t,\mu_3}$ and ϵ^+_{t,μ_4} obtained by fitting an AR model of order p. Insertion in (2.40) simplifies matters and we have to deal with the three summands in (2.41)-(2.43) separately. Consider first (2.42). Here, the sum over j_2 in (2.40) collapses and a rearrangement results in

$$
\begin{aligned}
&\frac{4\pi^2}{n}\sum_{j_1=-N}^{N}\sum_{m_1,m_3=1}^{r}\widetilde{q}_{h_1m_1}(\omega_{j_1})\left(\frac{1}{2\pi}\sum_{\mu_1=1}^{r}\left(\sum_{\nu_1=0}^{\infty}\left(\widehat{\mathbf{C}}_{\nu_1}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_1\mu_1}e^{-i\nu_1\omega_{j_1}}\right)\right.\\
&\times\left.\left(\overline{\sum_{\nu_3=0}^{\infty}\left(\widehat{\mathbf{C}}_{\nu_3}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_3\mu_1}e^{-i\nu_3\omega_{j_1}}}\right)^T\right)\overline{\widetilde{q}_{m_3k_1}(\omega_{j_1})}^T\\
&\times\sum_{m_2,m_4=1}^{r}\widetilde{q}_{k_2m_4}(\omega_{j_1})\left(\frac{1}{2\pi}\sum_{\mu_4=1}^{r}\left(\sum_{\nu_4=0}^{\infty}\left(\widehat{\mathbf{C}}_{\nu_4}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_4\mu_4}e^{-i\nu_4\omega_{j_1}}\right)\right.\\
&\times\left.\left(\overline{\sum_{\nu_2=0}^{\infty}\left(\widehat{\mathbf{C}}_{\nu_2}(p)\widehat{\mathbf{\Sigma}}^{1/2}(p)\right)_{m_2\mu_4}e^{-i\nu_2\omega_{j_1}}}\right)^T\right)\overline{\widetilde{q}_{m_2h_2}(\omega_{j_1})}^T e^{-i(k-h)\omega_{j_1}}\\
=\;&\frac{4\pi^2}{n}\sum_{j_1=-N}^{N}\left(\widetilde{\mathbf{Q}}(\omega_{j_1})\widehat{\mathbf{f}}_{AR}(\omega_{j_1})\overline{\widetilde{\mathbf{Q}}(\omega_{j_1})}^T\right)_{h_1k_1}\left(\widetilde{\mathbf{Q}}(\omega_{j_1})\widehat{\mathbf{f}}_{AR}(\omega_{j_1})\overline{\widetilde{\mathbf{Q}}(\omega_{j_1})}^T\right)_{k_2h_2}\\
&\times e^{-i(k-h)\omega_{j_1}}.
\end{aligned}
$$

Because of the uniform convergence in ω of the quantities $\widetilde{\mathbf{Q}}(\omega)$ and $\widehat{\mathbf{f}}_{AR}(\omega)$, the Riemann sum above converges to

$$
\begin{aligned}
&2\pi\int_{-\pi}^{\pi}\left(\mathbf{Q}(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T\right)_{h_1k_1}\left(\mathbf{Q}(\omega)\mathbf{f}_{AR}(\omega)\overline{\mathbf{Q}(\omega)}^T\right)_{k_2h_2}e^{-i(k-h)\omega}d\omega\\
=\;&2\pi\int_{-\pi}^{\pi}f_{h_1k_1}(\omega)f_{k_2h_2}(\omega)e^{-i(k-h)\omega}d\omega
\end{aligned}
$$

in probability. Finally, the multivariate inversion formula yields

$$
\begin{aligned}
& 2\pi \int_{-\pi}^{\pi} f_{h_1k_1}(\omega) f_{k_2h_2}(\omega) e^{-i(k-h)\omega} d\omega \\
= \; & 2\pi \int_{-\pi}^{\pi} f_{h_1k_1}(\omega) \frac{1}{2\pi} \sum_{t=-\infty}^{\infty} e^{-it\omega} \gamma_{k_2h_2}(t) e^{-i(k-h)\omega} d\omega \\
= \; & \sum_{t=-\infty}^{\infty} \gamma_{k_2h_2}(t) \int_{-\pi}^{\pi} f_{h_1k_1}(\omega) e^{i(-t-k+h)\omega} d\omega \\
= \; & \sum_{t=-\infty}^{\infty} \gamma_{h_2k_2}(-t) \gamma_{h_1k_1}(-t-k+h).
\end{aligned}
$$

Arguments are analogue for (2.43) and its limit in probability is

$$
\sum_{t=-\infty}^{\infty} \gamma_{h_2k_2}(-t) \gamma_{h_1k_1}(-t-k+h).
$$

It remains to check (2.41). Inserting in (2.40) and rearranging gives the expression

$$
\begin{aligned}
& \sum_{s_1,s_2,s_3,s_4=1}^{r} \left(\frac{2\pi}{n} \sum_{j_1=-N}^{N} \left(\widetilde{\mathbf{Q}}(\omega_{j_1}) \left(\frac{1}{\sqrt{2\pi}} \sum_{\nu_1=0}^{\infty} \widehat{\mathbf{C}}_{\nu_1}(p) e^{-i\nu_1\omega_{j_1}} \right) \right) \right)_{h_1s_1} \\
& \times \left(\left(\overline{\left(\frac{1}{\sqrt{2\pi}} \sum_{\nu_2=0}^{\infty} \widehat{\mathbf{C}}_{\nu_2}(p) e^{-i\nu_2\omega_{j_1}} \right)}^{T} \overline{\widetilde{\mathbf{Q}}(\omega_{j_1})}^{T} \right)_{s_2h_2} e^{ih\omega_{j_1}} \right) \\
& \times \sum_{\mu_1,\mu_2,\mu_3,\mu_4=1}^{r} \widehat{\mathbf{\Sigma}}^{1/2}(p)_{s_1\mu_1} \widehat{\mathbf{\Sigma}}^{1/2}(p)_{s_2\mu_2} \widehat{\mathbf{\Sigma}}^{1/2}(p)_{s_3\mu_3} \widehat{\mathbf{\Sigma}}^{1/2}(p)_{s_4\mu_4} \widetilde{\kappa}_4(p; \mu_1, \mu_2, \mu_3, \mu_4) \\
& \times \left(\frac{2\pi}{n} \sum_{j_2=-N}^{N} \left(\widetilde{\mathbf{Q}}(\omega_{j_2}) \left(\frac{1}{\sqrt{2\pi}} \sum_{\nu_3=0}^{\infty} \widehat{\mathbf{C}}_{\nu_3}(p) e^{-i\nu_3\omega_{j_2}} \right) \right) \right)_{k_1s_3} \\
& \times \left(\left(\overline{\left(\frac{1}{\sqrt{2\pi}} \sum_{\nu_4=0}^{\infty} \widehat{\mathbf{C}}_{\nu_4}(p) e^{-i\nu_4\omega_{j_2}} \right)}^{T} \overline{\widetilde{\mathbf{Q}}(\omega_{j_2})}^{T} \right)_{s_4k_2} e^{ik\omega_{j_2}} \right),
\end{aligned}
$$

which converges to the corresponding part as stated in the theorem. □

Proof of Theorem 2.5.4

As in the proof of Theorem 2.5.2, it is more convenient to show asymptotic normality for

$$
\mathcal{L}\left(\sqrt{nb}\, vec\left(\left[\widehat{\mathbf{\Gamma}}^*(0) - E^+[\widehat{\mathbf{\Gamma}}^*(0)] | \cdots | \widehat{\mathbf{\Gamma}}^*(s) - E^+[\widehat{\mathbf{\Gamma}}^*(s)] \right] \right) | \underline{X}_1, \ldots, \underline{X}_n \right)
$$

and analogue to the proof of Lemma 2.7.3 it suffices here to consider the case $s = 1$ with some lag h. Hence, we can focus on

$$\sqrt{n} bvec\left(\widehat{\mathbf{\Gamma}}^*(h) - E^+[\widehat{\mathbf{\Gamma}}^*(h)]\right).$$

Recall that $\widehat{\mathbf{\Gamma}}^*(h)$ can be displayed as a so-called spectral mean [cf. Dahlhaus (1985) for the univariate case], that is

$$\widehat{\mathbf{\Gamma}}^*(h) = \int_{\pi}^{\pi} \mathbf{I}_n^*(\omega) e^{ih\omega} d\omega.$$

Using $\mathbf{I}_n^*(\omega) = \widetilde{\mathbf{Q}}(\omega)\mathbf{I}_n^+(\omega)\overline{\widetilde{\mathbf{Q}}(\omega)}^T$ and $\mathbf{I}_n^+(\omega) = \frac{1}{2\pi}\sum_{k=-(n-1)}^{n-1} \widehat{\mathbf{\Gamma}}^+(k)e^{-ik\omega}$, where $\widehat{\mathbf{\Gamma}}^+(h)$ is analogue to (2.18) based on $\underline{X}_1^+, \ldots, \underline{X}_n^+$, we get

$$\widehat{\mathbf{\Gamma}}^*(h) = \sum_{k=-(n-1)}^{n-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \widetilde{\mathbf{Q}}(\omega)\widehat{\mathbf{\Gamma}}^+(k)\overline{\widetilde{\mathbf{Q}}(\omega)}^T e^{-i(k-h)\omega} d\omega.$$

Further, due to the formula $vec(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})vec(\mathbf{B})$ for appropriate matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, application of the vec-operator yields

$$\begin{aligned}
&\sqrt{n} vec(\widehat{\mathbf{\Gamma}}^*(h) - E^+[\widehat{\mathbf{\Gamma}}^*(h)]) \\
= &\sum_{k=-(n-1)}^{n-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\overline{\widetilde{\mathbf{Q}}(\omega)} \otimes \widetilde{\mathbf{Q}}(\omega)\right) e^{-i(k-h)\omega} d\omega \sqrt{n}\ vec\left(\widehat{\mathbf{\Gamma}}^+(k) - E^+[\widehat{\mathbf{\Gamma}}^+(k)]\right) \\
= &\underline{Z}_n^+.
\end{aligned}$$

To make Proposition 6.3.9 in Brockwell and Davis (1991) applicable, let $M \in \mathbb{N}$ be fixed and split the last sum in two parts to obtain

$$\begin{aligned}
&\sum_{k=-M}^{M} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\overline{\widetilde{\mathbf{Q}}(\omega)} \otimes \widetilde{\mathbf{Q}}(\omega)\right) e^{-i(k-h)\omega} d\omega \sqrt{n}\ vec\left(\widehat{\mathbf{\Gamma}}^+(k) - E^+[\widehat{\mathbf{\Gamma}}^+(k)]\right) \\
&+ \sum_{\substack{k=-(n-1) \\ |k|>M}}^{n-1} \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\overline{\widetilde{\mathbf{Q}}(\omega)} \otimes \widetilde{\mathbf{Q}}(\omega)\right) e^{-i(k-h)\omega} d\omega \sqrt{n}\ vec\left(\widehat{\mathbf{\Gamma}}^+(k) - E^+[\widehat{\mathbf{\Gamma}}^+(k)]\right), \\
= &\underline{Z}_{n,M}^+ + (\underline{Z}_n^+ - \underline{Z}_{n,M}^+),
\end{aligned}$$

with an obvious notation for $\underline{Z}_{n,M}^+$. Now, it suffices to have that for all $M \in \mathbb{N}$ the quantity $\underline{Z}_{n,M}^+$ converges weakly to a normal distribution in probability depending on M, which itself in turn converges for $M \to \infty$. Moreover, for all $\epsilon > 0$, the condition

$$\lim_{M\to\infty} \limsup_{n\to\infty} P^+(\|\underline{Z}_n - \underline{Z}_{n,M}\| > \epsilon) = 0 \tag{2.44}$$

in probability has to be satisfied. At first, let M be fixed. Then $\underline{Z}^+_{n,M}$ can be displayed as a matrix-vector product and we get

$$\begin{aligned}\underline{Z}^+_{n,M} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left[\left(\overline{\widetilde{\mathbf{Q}}(\omega)}\otimes\widetilde{\mathbf{Q}}(\omega)\right)e^{-i(-M-h)\omega}\right|\cdots\left|\left(\overline{\widetilde{\mathbf{Q}}(\omega)}\otimes\widetilde{\mathbf{Q}}(\omega)\right)e^{-i(M-h)\omega}\right]d\omega \\ &\quad\cdot\sqrt{n}vec\left(\left[\widehat{\mathbf{\Gamma}}^+(-M)-E^+[\widehat{\mathbf{\Gamma}}^+(-M)]\right|\cdots\left|\widehat{\mathbf{\Gamma}}^+(M)-E^+[\widehat{\mathbf{\Gamma}}^+(M)]\right]\right) \\ &= \mathbf{H}^+_{n,M}\cdot\underline{R}^+_{n,M},\end{aligned}$$

where the $(r^2\times(2n-1)r^2)$-matrix $\mathbf{H}^+_{n,M}$ is multiplied with the $(2n-1)r^2$-dimensional vector $\underline{R}^+_{n,M}$. Applying an adequate CLT [e.g. the CLT in Neumann and Paparoditis (2008)], we get asymptotic normality of $\underline{R}^+_{n,M}$, which contains nothing else but empirical autocovariances of the usual residual AR-bootstrap. Together with the convergence in probability of $\mathbf{H}^+_{n,M}$ and Slutsky we get the required weak convergence of $\underline{Z}^+_{n,M}=\mathbf{H}^+_{n,M}\cdot\underline{R}^+_{n,M}$ and its asymptotic multivariate normal distribution depending on M converges itself to the correct covariance matrix as $M\to\infty$ by Theorem 2.5.3. It remains to show (2.44). By Markov inequality, it suffices to consider

$$\begin{aligned}&E^+[\|\underline{Z}_n-\underline{Z}_{n,M}\|] \\ &= \|\sum_{\substack{k=-(n-1)\\|k|>M}}^{n-1}\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\overline{\widetilde{\mathbf{Q}}(\omega)}\otimes\widetilde{\mathbf{Q}}(\omega)\right)e^{-i(k-h)\omega}d\omega\sqrt{n}vec\left(\widehat{\mathbf{\Gamma}}^+(k)-E^+[\widehat{\mathbf{\Gamma}}^+(k)]\right)\| \\ &\le \sum_{\substack{k=-(n-1)\\|k|>M}}^{n-1}\|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\overline{\widetilde{\mathbf{Q}}(\omega)}\otimes\widetilde{\mathbf{Q}}(\omega)\right)e^{-i(k-h)\omega}d\omega\|\|\sqrt{n}vec\left(\widehat{\mathbf{\Gamma}}^+(k)-E^+[\widehat{\mathbf{\Gamma}}^+(k)]\right)\| \\ &= O_P(1)\sum_{\substack{k=-(n-1)\\|k|>M}}^{n-1}\|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\overline{\widetilde{\mathbf{Q}}(\omega)}\otimes\widetilde{\mathbf{Q}}(\omega)\right)e^{-i(k-h)\omega}d\omega\|, \qquad (2.45)\end{aligned}$$

where we have used $\sqrt{n}\ vec(\widehat{\mathbf{\Gamma}}^+(k)-E^+[\widehat{\mathbf{\Gamma}}^+(k)])=O_P(1)$ uniformly in k. Now, let $\|\cdot\|$ be the 1-norm for matrices $\|\cdot\|_1$, defined as $\|\mathbf{A}\|_1=\sum_{i,j}|a_{i,j}|$. The normed expression in (2.45) is a matrix, whose entries are usual Fourier coefficients of the type

$$a_{k-h}(r,s,t,u)=\frac{1}{2\pi}\int_{-\pi}^{\pi}\overline{\widetilde{q}_{rs}(\omega)}\widetilde{q}_{tu}(\omega)e^{-i(k-h)\omega}d\omega,$$

where $r,s,t,u\in\{1,\dots,r\}$. Because of Lemma 2.7.1, the function $\overline{\widetilde{q}_{rs}(\cdot)}\widetilde{q}_{tu}(\cdot)$ is three times differentiable and therefore $|a_{k-h}|$ can be bounded by $\frac{T_n}{|k-h|^2}$, where

$$T_n=\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{\partial^2}{\partial\omega^2}\left(\overline{\widetilde{q}_{rs}(\omega)}\widetilde{q}_{tu}(\omega)\right)\right|d\omega=O_P(1)$$

uniformly in k. Finally, for sufficiently large M, we obtain

$$\lim_{M\to\infty}\limsup_{n\to\infty} E^+[\|\underline{Z}_n - \underline{Z}_{n,M}\|] \leq \lim_{M\to\infty}\limsup_{n\to\infty} O_P(1) \sum_{\substack{k=-(n-1)\\ |k|>M}}^{n-1} \frac{1}{|k-h|^2} = 0$$

in probability, where we have used $\sum_{|k|>M} \frac{1}{|k-h|^2} < \infty$. This concludes the proof. □

Proof of Corollary 2.5.4

All we have to show is that the asymptotic covariance derived in Theorem 2.5.3 agrees with (2.20). It suffices to consider the first part containing the fourth order cumulants, because the second parts are already equal. Note, that the sums over ν_1 and ν_2 in (2.19) are from 0 to ∞ due to causality. Under the assumption of an underlying $VAR(p_0)$-model, fitting a $VAR(p)$-model with $p \geq p_0$, we estimate the parameters $\mathbf{A}_1, \ldots, \mathbf{A}_{p_0}$ consistently with $\widehat{\mathbf{A}}_1(p), \ldots, \widehat{\mathbf{A}}_p(p)$, where $\widehat{\mathbf{A}}_k(p)$ converges to zero for $k > p_0$. Thus, we obtain $\mathbf{C}_\nu(p) = \mathbf{C}_\nu$ for all $\nu \in \mathbb{N}_0$ on the one hand and $\mathbf{f} = \mathbf{f}_{AR}$ on the other hand, which in turn yields $\widetilde{\mathbf{Q}}(\omega) \to \mathbf{Q}(\omega) = \mathbf{I}_r$ in probability uniformly in ω. Moreover, it holds

$$\kappa_4(s_1, s_2, s_3, s_4) = \kappa_4(p; s_1, s_2, s_3, s_4).$$

Together, the first part of expression (2.21) becomes

$$\sum_{s_1,s_2,s_3,s_4=1}^{r} \left(\int_{-\pi}^{\pi} (\mathbf{C}(\omega_1))_{js_1} \left(\overline{\mathbf{C}(\omega_1)}^T\right)_{s_2k} e^{ig\omega_1} d\omega_1\right)$$
$$\kappa_4(s_1, s_2, s_3, s_4) \left(\int_{-\pi}^{\pi} (\mathbf{C}(\omega_2))_{ls_3} \left(\overline{\mathbf{C}(\omega_2)}^T\right)_{s_4m} e^{ih\omega_2} d\omega_2\right),$$

which concludes the proof. □

Chapter 3

A new frequency domain approach of testing for covariance stationarity and for periodic stationarity in multivariate linear processes

Based on: Carsten Jentsch
A new frequency domain approach of testing for covariance stationarity and for periodic stationarity in multivariate linear processes.
Submitted (2010).

Abstract. In modeling seasonal time series data, periodically (non-)stationary processes have become quite popular over the last years and it is well known that these models may be represented as higher-dimensional stationary models. In this chapter, it is shown that the spectral density matrix of this higher-dimensional process exhibits a certain structure if and only if the observed process is covariance stationary. By exploiting this relationship, a new L_2-type test statistic is proposed for testing whether a multivariate periodically stationary linear process is even covariance stationary. Moreover, it is shown that this test may also be used to test for periodic stationarity. The asymptotic normal distribution of the test statistic under the null is derived and the test is shown to have an omnibus property. The chapter concludes with a simulation study, where the small sample performance of the test procedure is improved by using the hybrid bootstrap scheme.

2000 *Mathematics Subject Classification.* 62M10, 62G10.

Keywords and phrases. hypothesis testing; testing for stationarity; testing for periodic stationarity; periodic time series; multivariate time series; linear process; spectral density matrix; kernel spectral density estimates; hybrid bootstrap.

3.1 Introduction

In modelling time series the challenge is to find parsimonious models that satisfactorily capture the possibly complicated dependence structure of the observed data. The aim of parsimony becomes even more important for multivariate time series, where the number of involved parameters that may have to be estimated increases dramatically. Usually, second-order stationarity (covariance stationarity) is assumed to ensure sufficient mathematical tractability of the chosen model and to make things manageable. A common procedure in modelling non-stationary time series is to standardize or to filter the observed series and then fit an appropriate stationary stochastic model to the reduced series.

However, in many situations, the assumption of stationarity is not fulfilled and there is no transformation that may be applied to the data to achieve second-order stationarity. Sometimes and particularly for seasonal time series, this is because the covariance structure of a time series possibly depends on the season, that is, the autocovariance function is periodic for all lags. For instance, time series of this kind appear in hydrology, climatology, metereology and other geophysical sciences, but also in economics, where the observed time series are characterized by periodic variations in both the mean and covariance structure.

For time series with periodic correlation structure, Gladyshev (1961) introduced the notion of periodically stationary processes. Further pioneering work has been done by Jones and Brelsford (1967), Pagano (1978) and Troutman (1979), who have examined fundamental properties of univariate periodic autoregressive (PAR) processes. Later periodic moving-average processes (PMA) [see Cipra (1985), Bentarzi and Hallin (1994, 1998)] and the more general class of periodic autoregressive moving-average time series (PARMA) have been considered [cf. Vecchia (1985a,b), Lund and Basawa (2000), Basawa and Lund (2001) among others]. These models are extensions of the usual ARMA models where the coefficients and the variances of the white noise process are allowed to depend on the season. Multivariate generalizations of these models have been investigated by Ula (1990, 1993), Franses and Paap (2004) and Lütkepohl (2005), but basic research still has to be done.

Time series analysis of data sequences usually involves three main steps: model identification, parameter estimation and diagnostic checking. Concerned with model identification in seasonal time series, it seems natural to decide first whether there are actually periodicities present in the data and if so to determine the period, that is, the smallest integer s, so that all autocovariances are periodic with s periods, and to choose the orders of a potential PARMA(p,q) model afterwards.

Since a periodic autocorrelation structure complicates all three steps of model building extensively and the number of parameters in a nonstationary periodic model

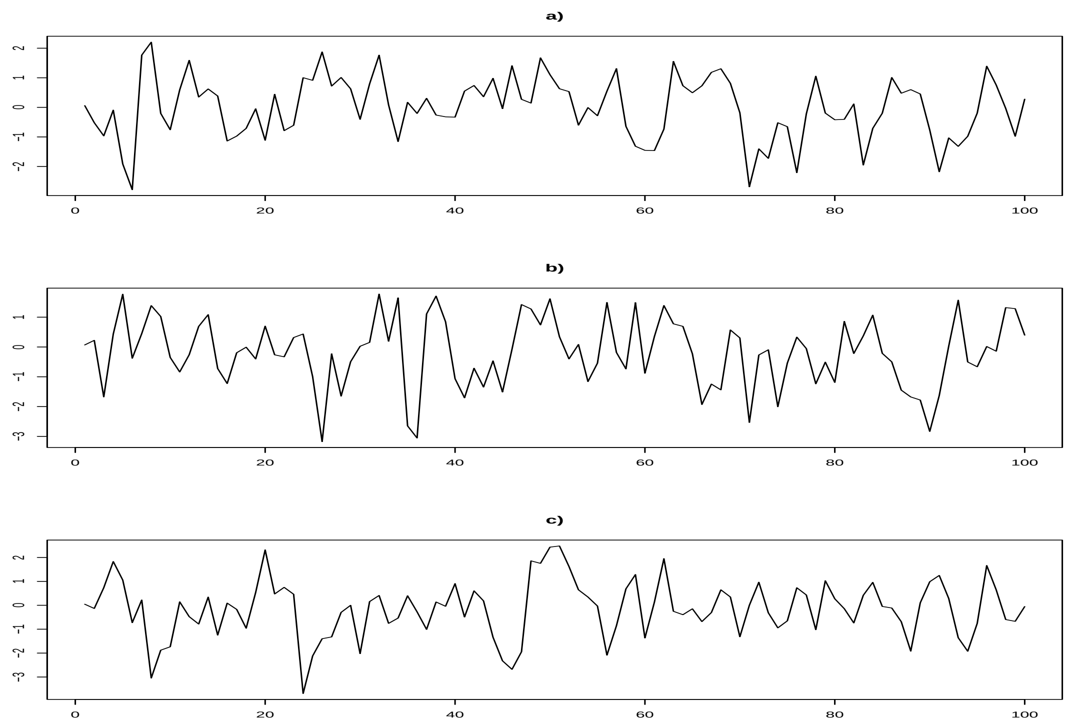

Figure 3.1: Realizations of length $N = 100$ of a) MA(1) process with $b_1 = 0.5$ and i.i.d $\mathcal{N}(0,1)$ white noise, b) periodic MA(1) process with $s = 2$ periods, $b_1^{(1)} = 0.3$, $b_1^{(2)} = 0.7$ and i.i.d $\mathcal{N}(0,1)$ white noise, c) periodic MA(1) process with $s = 2$ periods, $b_1^{(1)} = b_1^{(2)} = 0.5$ and independent $\mathcal{N}(0,\sigma_j^2)$ white noise with $\sigma_1 = 0.8$ and $\sigma_2 = 1.2$.

involves a possible s fold increase in the number of parameters over that in a stationary nonperiodic model, care must be exercised in its application. For instance, Lund, Shao and Basawa (2006) have investigated parsimonious representations in the class of periodic time series models, in order to reduce the number of parameters.

The consequences of fitting a stationary AR model to data generated by a nonstationary periodic AR model have been emphasized in Tiao and Grupe (1980). In this case the resulting model is misspecified which causes a considerable forecasting error as shown in their paper. Regarding Figure 3.1, it is usually not possible to distinguish between stationary and periodically stationary processes or to guess the number of periods from the plot of the data. Therefore, it is of large interest to establish powerful tools to decide whether the assumptions of covariance stationarity may be valid for a particular (multiple) seasonal time series or to determine its actual period, respectively.

Cipra (1985) developed a Durbin-type estimation procedure for the coefficients of

PMA models and suggested a test for periodic structure based on corresponding asymptotic normality results. Vecchia and Ballerini (1991) proposed a test for deciding whether periodicities exist in the autocorrelation function of a seasonal time series under the assumptions of a causal periodic linear model. Their approach is based on a Fourier-transformed version of the estimated periodic autocorrelation function. In McLeod (1993, 1994) portmanteau type test statistics based on residuals of a fitted AR(I)MA model have been used and the three usual stages of model building methodology have been illustrated in detail. The graphical approach of Hurd and Gerr (1991) is based on the spectral representation for harmonizable second-order sequences and they suggest to use Goodman's coherence statistic to test for periodic correlation. Lenart, Leśkow and Synowiecki (2008) proposed a test statistic that exploits properties of the Fourier coefficients of the time-dependent autocovariance function for univariate periodically correlated time series under mixing assumptions using subsampling. Recently, Ursu and Duchesne (2009) considered vector-valued periodic autoregressive models (PVAR) and developed multivariate generalizations of theorems concerned with portmanteau-type tests obtained by McLeod (1994).

So far, most of the present literature on periodically stationary models concentrates on univariate periodic time series and/or finite parametric models as PAR and PARMA. However, multivariate models are expected to be extremely useful in practice and powerful tools for detecting periodicities in general linear models are of considerable interest.

Hence, this chapter deals with the very general class of vector-valued periodically stationary linear time series (PVL) defined in (3.1) below. Observe that PVL models include the important subclasses of multivariate PAR and multivariate PARMA models.

It is well known that any d-dimensional periodically stationary process with s periods may be expressed as an sd-dimensional stationary process as shown in (3.3) for linear processes. Usually, this technique is associated with Gladyshev [cf. Gladyshev (1961)]. For instance, Pagano (1978) among others considered the relation of periodic and multiple autoregression in the univariate situation. This relationship allows using the classical theory of stationary time series as, for instance, nonparametric estimation in the frequency domain. Many relevant hypothesis about second-order properties of multivariate stationary time series may be expressed in terms of the spectral density matrix and the formulation of hypothesis in the frequency domain enjoys the advantage of a general nonparametric framework based on, for example, kernel spectral density estimators. In the context of periodically stationary processes, properties of the spectral density matrix have been considered by Troutman (1979) and Sakai (1991).

The main purpose of this chapter is to present a test procedure for deciding whether the underlying PVL process with a (predetermined) number of s periods actually is covariance stationary. But, additionally, it is shown that this test may be applied also for testing whether the underlying PVL time series is periodically stationary with some period smaller than s, which to the author's knowledge has not been investigated yet thoroughly in statistical literature. An L_2-type test statistic that estimates the integrated deviation from the null hypothesis is suggested that exploits the specific shape of a slightly adjusted spectral density matrix of the corresponding higher-dimensional process under the null hypothesis.

A CLT for the test statistic is proved and the test is shown to be an omnibus test that has power against any alternative. The finite sample performance of this test is checked in a simulation study using critical values obtained from the CLT and from an appropriate bootstrap procedure. Here, we make use of the multiple hybrid bootstrap proposed by Jentsch and Kreiss (2010), which is well suited for kernel spectral density estimation in the situation of multivariate linear processes. The use of bootstrap methods for testing periodicities in (autoregressive) time series models was already recommended by Herwartz (1998).

The chapter is organized as follows. In Section 3.2, some preliminary results concerning d-variate periodically stationary linear processes are presented and some examples are discussed. In particular, Theorem 3.2.1 provides the specific structure of autocovariance function and (modified) spectral density of the corresponding sd-variate process under the null hypothesis. Section 3.3 deals with the construction of the test statistic, its asymptotic normality in Theorem 3.3.1 and its omnibus property in Theorem 3.3.2. A small simulation study is presented in Section 3.4. Finally, proofs of the main results can be found in Section 3.5.

3.2 Preliminary results

Let $(\underline{Y}_t, t \in \mathbb{Z})$ be a d-dimensional periodically stationary linear process with s periods, $s \in \mathbb{N}$, that is,

$$\underline{Y}_{j+sT} = \sum_{k=-\infty}^{\infty} \mathbf{b}_k^{(j)} \underline{\epsilon}_{j+sT-k}, \quad j = 1, \ldots, s, \quad T \in \mathbb{Z}, \tag{3.1}$$

where $(\underline{\epsilon}_t, t \in \mathbb{Z})$ are independent d-variate random variables with mean $E(\underline{\epsilon}_t) = 0$ for all $t \in \mathbb{Z}$, covariance matrices $E(\underline{\epsilon}_{j+sT}\underline{\epsilon}_{j+sT}^T) = \mathbf{\Sigma}_j$ for $j = 1, \ldots, s$ and all $T \in \mathbb{Z}$. Further, for $j = 1, \ldots, s$, $\mathbf{b}_k^{(j)}$, $k \in \mathbb{Z}$ are $(d \times d)$ coefficient matrices and $\mathbf{b}_0^{(j)} = \mathbf{I}_d$ are $(d \times d)$ unit matrices.

Note that the process in (3.1) is supposed to have zero mean, that is $E(\underline{Y}_t) = 0$. Therefore, in practice, it is assumed that the analysis of the mean has been done, that is, the possibly periodic mean is zero or has been removed, and we are concerned here only with methods for determining whether the autocovariances have periodic structure.

In general, a periodically stationary process is *not* covariance stationary, because its $(d \times d)$ autocovariance function is periodic with period s for all lags $h \in \mathbb{Z}$. More precisely, using the notation

$$\mathbf{\Gamma}_Y(h,m) = Cov(\underline{Y}_{m+sT}, \underline{Y}_{m+sT-h}) = E[\underline{Y}_{m+sT}\underline{Y}^T_{m+sT-h}], \tag{3.2}$$

it holds $\mathbf{\Gamma}_Y(h,m) = \mathbf{\Gamma}_Y(h,m+sk)$ and $\mathbf{\Gamma}_Y(-h,m) = \mathbf{\Gamma}_Y^T(h,m+h)$ for all $m = 1,\dots,s$, $k \in \mathbb{Z}$ and all lags $h \in \mathbb{Z}$.

In the theory of periodically stationary time series, it is a common technique to interpret them as higher-dimensional stationary processes. Precisely, the d-variate process $(\underline{Y}_t, t \in \mathbb{Z})$ defined in (3.1) may be represented as an sd-dimensional covariance stationary process $(\underline{X}_t, t \in \mathbb{Z})$ according to

$$\begin{aligned}\underline{X}_T = \begin{pmatrix} \underline{Y}_{1+sT} \\ \underline{Y}_{2+sT} \\ \vdots \\ \underline{Y}_{s+sT} \end{pmatrix} &= \sum_{k=-\infty}^{\infty} \begin{pmatrix} \mathbf{b}^{(1)}_{sk} & \mathbf{b}^{(1)}_{sk-1} & \cdots & \mathbf{b}^{(1)}_{sk-(s-1)} \\ \mathbf{b}^{(2)}_{sk+1} & \mathbf{b}^{(2)}_{sk} & & \mathbf{b}^{(2)}_{sk-(s-2)} \\ \vdots & & \ddots & \vdots \\ \mathbf{b}^{(s)}_{sk+(s-1)} & \mathbf{b}^{(s)}_{sk+(s-2)} & \cdots & \mathbf{b}^{(s)}_{sk} \end{pmatrix} \begin{pmatrix} \underline{\epsilon}_{1+s(T-k)} \\ \underline{\epsilon}_{2+s(T-k)} \\ \vdots \\ \underline{\epsilon}_{s+s(T-k)} \end{pmatrix} \\ &= \sum_{k=-\infty}^{\infty} \mathbf{B}_k \underline{e}_{T-k}, \quad T \in \mathbb{Z} \end{aligned} \tag{3.3}$$

with an obvious notation for the $(sd \times sd)$ matrices $\mathbf{B}_k$, $k \in \mathbb{Z}$ and the sd-dimensional white noise process $(\underline{e}_t, t \in \mathbb{Z})$. Further, it holds $E(\underline{e}_t) = \underline{0}$ and $E(\underline{e}_t\underline{e}_t^T) = \mathbf{\Sigma}_e$ with block-diagonal $(sd \times sd)$ covariance matrix $\mathbf{\Sigma}_e = diag(\mathbf{\Sigma}_1, \mathbf{\Sigma}_2, \dots, \mathbf{\Sigma}_s)$.

Observe that due to the algebraic equivalence between multivariate stationarity of an sd-variate process and periodic stationarity with s periods of a d-variate process, the process $(\underline{X}_t, t \in \mathbb{Z})$ introduced in (3.3) is stationary if and only if $(\underline{Y}_t, t \in \mathbb{Z})$ defined in (3.1) is periodically stationary [cf. Gladyshev (1961), Ula (1990) and Ursu and Duchesne (2009)]. Moreover, note that covariance stationarity always implies periodic stationarity with any number of periods $s \geq 2$ and periodic stationarity with s periods implies periodic stationary with ks periods for any $k \in \mathbb{N}$.

Under sufficient summability assumptions on $(\mathbf{B}_k, k \in \mathbb{Z})$, the process $(\underline{X}_t, t \in \mathbb{Z})$ is covariance stationary. Therefore, its $(sd \times sd)$ autocovariance function $\mathbf{\Gamma}(h)$,

$$\mathbf{\Gamma}(h) = \sum_{k=-\infty}^{\infty} \mathbf{B}_{k+h}\mathbf{\Sigma}_e\mathbf{B}_k^T = \begin{pmatrix} \mathbf{\Gamma}_{11}(h) & \mathbf{\Gamma}_{12}(h) & \cdots & \mathbf{\Gamma}_{1s}(h) \\ \mathbf{\Gamma}_{21}(h) & \mathbf{\Gamma}_{22}(h) & & \mathbf{\Gamma}_{2s}(h) \\ \vdots & & \ddots & \vdots \\ \mathbf{\Gamma}_{s1}(h) & \mathbf{\Gamma}_{s2}(h) & \cdots & \mathbf{\Gamma}_{ss}(h) \end{pmatrix}, \quad h \in \mathbb{Z}, \tag{3.4}$$

exists, where $\mathbf{\Gamma}_{mn}(h)$ are $(d \times d)$ block matrices with $\mathbf{\Gamma}_{mn}(-h) = \mathbf{\Gamma}_{mn}^T(h)$ for all m, n. Also, under suitable assumptions, $(\underline{X}_t, t \in \mathbb{Z})$ exhibits an $(sd \times sd)$ spectral density matrix $\mathbf{f}(\omega)$, which has the representation

$$\begin{aligned}\mathbf{f}(\omega) &= \frac{1}{2\pi}\left(\sum_{k=-\infty}^{\infty} \mathbf{B}_k e^{-ik\omega}\right) \mathbf{\Sigma}_e \left(\sum_{k=-\infty}^{\infty} \mathbf{B}_k e^{-ik\omega}\right)^H \\ &= \begin{pmatrix} \mathbf{F}_{11}(\omega) & \mathbf{F}_{12}(\omega) & \cdots & \mathbf{F}_{1s}(\omega) \\ \mathbf{F}_{21}(\omega) & \mathbf{F}_{22}(\omega) & & \mathbf{F}_{2s}(\omega) \\ \vdots & & \ddots & \vdots \\ \mathbf{F}_{s1}(\omega) & \mathbf{F}_{s2}(\omega) & \cdots & \mathbf{F}_{ss}(\omega) \end{pmatrix}, \quad \omega \in [-\pi, \pi], \end{aligned} \tag{3.5}$$

where $\mathbf{F}_{mn}(\omega)$ are $(d \times d)$ block matrices with $\mathbf{F}_{mn}(\omega) = \mathbf{F}_{nm}^H(\omega)$ for all m, n. Here and throughout the chapter, all matrix-valued quantities are written as bold letters, all vector-valued quantities are underlined and $(\mathbf{A}^H)$ $\mathbf{A}^T$ indicates the (conjugate) transpose of a matrix $\mathbf{A}$. Moreover, minuscules $\mathbf{f}(\omega)$ are $(sd \times sd)$ and capitals $\mathbf{F}_{mn}(\omega)$ are $(d \times d)$ matrices unless otherwise stated.

In this chapter, we are mainly concerned with statistical procedures for testing whether the observed process $(\underline{Y}_t, t \in \mathbb{Z})$ is covariance stationary. Hence, the question arises whether the spectral density matrix $\mathbf{f}(\omega)$ of $(\underline{X}_t, t \in \mathbb{Z})$ defined in (3.5) has some specific and possibly unique shape if the underlying process $(\underline{Y}_t, t \in \mathbb{Z})$ is indeed second order stationary. To answer the question just posed, it seems convenient not to deal with $\mathbf{f}(\omega)$ itself, but with some apropriately adjusted quantity $\mathbf{g}(\omega)$ defined below.

Let $\mathbf{d}(\omega) = diag(\mathbf{D}_{1,s}(\omega), \ldots, \mathbf{D}_{s,s}(\omega))$ be an $(sd \times sd)$ diagonal matrix, where $\mathbf{D}_{j,s}(\omega) = e^{-i\frac{j}{s}\omega}\mathbf{I}_d$, $j = 1, \ldots, s$ are $(d \times d)$ diagonal matrices. Now, define $\mathbf{g}(\omega)$ according to

$$\mathbf{g}(\omega) = \mathbf{d}(\omega)\mathbf{f}(\omega)\mathbf{d}^H(\omega), \quad \omega \in [-\pi, \pi]. \tag{3.6}$$

The $(sd \times sd)$-valued function $\mathbf{g}(\omega)$ is called *modified spectral density* of $(\underline{X}_t, t \in \mathbb{Z})$ from now on. This quantity has already been used by Troutman (1979) to describe the limiting spectral density of a nonstationary periodic autoregressive model. Similar to equation (3.5), $(d \times d)$ matrices $\mathbf{G}_{mn}(\omega)$ are introduced by

$$\mathbf{g}(\omega) = \begin{pmatrix} \mathbf{G}_{11}(\omega) & \mathbf{G}_{12}(\omega) & \cdots & \mathbf{G}_{1s}(\omega) \\ \mathbf{G}_{21}(\omega) & \mathbf{G}_{22}(\omega) & & \mathbf{G}_{2s}(\omega) \\ \vdots & & \ddots & \vdots \\ \mathbf{G}_{s1}(\omega) & \mathbf{G}_{s2}(\omega) & \cdots & \mathbf{G}_{ss}(\omega) \end{pmatrix} \tag{3.7}$$

and due to Hermitianity of $\mathbf{f}(\omega)$, this property holds for $\mathbf{g}(\omega)$ as well, that is, $\mathbf{G}_{mn}(\omega) = \mathbf{G}_{nm}^H(\omega)$ for all m, n.

The following Theorem 3.2.1 provides exactly the relationship between covariance stationarity of $(\underline{Y}_t, t \in \mathbb{Z})$, the autocovariance function $\mathbf{\Gamma}(h)$ of $(\underline{X}_t, t \in \mathbb{Z})$ and the modified spectral density $\mathbf{g}(\omega)$ of $(\underline{X}_t, t \in \mathbb{Z})$.

Theorem 3.2.1 (Properties of $(\underline{X}_t, t \in \mathbb{Z})$ under stationarity of $(\underline{Y}_t, t \in \mathbb{Z})$). *Let $s \geq 2$. Under assumption (A) in Section 3.3, the following assertions* (i), (ii) *and* (iii) *are equivalent:*

(i) The d-variate process $(\underline{Y}_t, t \in \mathbb{Z})$ in (3.1) *is covariance stationary.*

(ii) The autocovariance function $\mathbf{\Gamma}(h)$ of the sd-variate process $(\underline{X}_t, t \in \mathbb{Z})$ in (3.4) *fulfills*

(iia) $\mathbf{\Gamma}_{m,m+r}(h) = \mathbf{\Gamma}_{n,n+r}(h)$ *and* $\mathbf{\Gamma}_{m+r,m}(h) = \mathbf{\Gamma}_{n+r,n}(h)$ *for all* $r = 0, \ldots, s-1$, $m, n = 1, \ldots, s-r$ *and* $h \in \mathbb{Z}$,

(iib) $\mathbf{\Gamma}_{m+r,m}(h) = \mathbf{\Gamma}_{n,n+s-r}(h+1)$ *for all* $r = 1, \ldots, s-1$, $m = 1, \ldots, s-r$, $n = 1, \ldots, r$ *and* $h \in \mathbb{Z}$.

(iii) The modified spectral density $\mathbf{g}(\omega)$ of the sd-variate process $(\underline{X}_t, t \in \mathbb{Z})$ in (3.7) *fulfills*

(iiia) $\mathbf{G}_{m,m+r}(\omega) = \mathbf{G}_{n,n+r}(\omega)$ *and* $\mathbf{G}_{m+r,m}(\omega) = \mathbf{G}_{n+r,n}(\omega)$ *for all* $r = 0, \ldots, s-1$, $m, n = 1, \ldots, s-r$ *and* $\omega \in [-\pi, \pi]$,

(iiib) $\mathbf{G}_{m,m+r}(\omega) = \mathbf{G}^H_{n,n+s-r}(\omega)$ *and* $\mathbf{G}_{m+r,m}(\omega) = \mathbf{G}^H_{n+s-r,n}(\omega)$ *for all* $r = 0, \ldots, s-1$, $m = 1, \ldots, s-r$, $n = 1, \ldots, r$ *and* $\omega \in [-\pi, \pi]$.

In principle, it is possible to use either the specific structure of the autocovariance function $\mathbf{\Gamma}(h)$ or of the modified spectral density $\mathbf{g}(\omega)$ to derive statistical procedures for testing whether the underlying process $(\underline{Y}_t, t \in \mathbb{Z})$ is covariance stationary. However, the focus in Section 3.3 of this chapter is on a test statistic based on the modified spectral density. For instance, Paparoditis (2000) discussed the advantages of frequency-domain tests based on spectral densities in comparison to time-domain tests based on autocovariances.

Subsequently, Remark 3.2.1 illustrates the specific shape of $\mathbf{g}(\omega)$ under stationarity of $(\underline{Y}_t, t \in \mathbb{Z})$ as stated in Theorem 3.2.1.

Remark 3.2.1 (On $\mathbf{g}(\omega)$ under covariance stationarity of $(\underline{Y}_t, t \in \mathbb{Z})$).

(i) Assertion (iiia) of Theorem 3.2.1 indicates that all block matrices on the principal diagonal of $\mathbf{g}(\omega)$ are equal. Moreover, all block matrices on one and the same (lower or upper) secondary diagonal of $\mathbf{g}(\omega)$ are equal.

(ii) Theorem 3.2.1 (iiib) establishes a relationship between the block matrices of distinct secondary diagonals of $\mathbf{g}(\omega)$*. Precisely, the block matrices on the* r*th upper (lower) diagonal are equal to the conjugate transpose matrices on the* $(s-r)$*th upper (lower) diagonal, when the principal diagonal is understood as* 0*th diagonal.*

(iii) Note that, for s *even, result (iiib) of Theorem 3.2.1 yields, that all* $(d \times d)$ *block matrices in the* $\frac{s}{2}$*th secondary diagonal of* $\mathbf{g}(\omega)$ *are hermitian itself. In the case of an univariate underlying time series* $(Y_t, t \in \mathbb{Z})$*, that is* $d = 1$*, this means all entries on the* $\frac{s}{2}$*th secondary diagonal are real-valued.*

To simplify notational issues concerning the indices in part (iii) of Theorem 3.2.1, introduce a more convenient modulo notation that is used throughout the chapter from now on. Define

$$\langle n \rangle = (n-1) mod\ s + 1 \in \{1, \ldots, s\}, \quad n \in \mathbb{Z}.$$

This convention is employed in Corollary 3.2.1 to unify notation and to derive a set of equations that is equivalent to assertion (iii) of Theorem 3.2.1.

Corollary 3.2.1 (Properties of $\mathbf{g}(\omega)$ under stationarity of $(\underline{Y}_t, t \in \mathbb{Z})$)**.**
Let $s \geq 2$*. Under the assumptions of Theorem 3.2.1, assertion (iiia) and (iiib) together are equivalent to*

$$\mathbf{G}_{mn}(\omega) = \mathbf{G}_{\langle m+j \rangle, \langle n+j \rangle}(\omega) \tag{3.8}$$

for all $m, n = 1, \ldots, s$*,* $j = 0, \ldots, s-1$ *and all* $\omega \in [-\pi, \pi]$*. In the following, we write briefly* $\mathbf{G}_{m+j,n+j}(\omega)$ *meaning* $\mathbf{G}_{\langle m+j \rangle, \langle n+j \rangle}(\omega)$*.*

The following Example 3.2.1 illustrates the specific shape of the modified spectral density matrix $\mathbf{g}(\omega)$ of process $(\underline{X}_t, t \in \mathbb{Z})$ in the situation of an underlying stationary d-variate $MA(1)$ model $(\underline{Y}_t, t \in \mathbb{Z})$ for different number of periods s.

Example 3.2.1 ($\mathbf{g}(\omega)$ for $MA(1)$ process $(\underline{Y}_t, t \in \mathbb{Z})$)**.**
Let $(\underline{Y}_t, t \in \mathbb{Z})$ *be a stationary* d*-variate* $MA(1)$ *process, that is,*

$$\underline{Y}_t = \underline{\epsilon}_t + \mathbf{b}\underline{\epsilon}_{t-1}, \quad t \in \mathbb{Z},$$

where $\underline{\epsilon}_t \sim (0, \Sigma)$ *is a* d*-variate i.i.d. white noise. The corresponding process* $(\underline{X}_t, t \in \mathbb{Z})$ *as defined in* (3.3) *becomes an* sd*-variate* $MA(1)$ *process and computing its mod-*

ified spectral density $\mathbf{g}_s(\omega)$ for $s = 2, 3, 4$ results in

$$\mathbf{g}_2(\omega) = \frac{1}{2\pi}\begin{pmatrix} \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T & \mathbf{\Sigma}\mathbf{b}^T e^{i\frac{1}{2}\omega} + \mathbf{b}\mathbf{\Sigma}e^{-i\frac{1}{2}\omega} \\ \mathbf{b}\mathbf{\Sigma}e^{-i\frac{1}{2}\omega} + \mathbf{\Sigma}\mathbf{b}^T e^{i\frac{1}{2}\omega} & \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T \end{pmatrix},$$

$$\mathbf{g}_3(\omega) = \frac{1}{2\pi}\begin{pmatrix} \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T & \mathbf{\Sigma}\mathbf{b}^T e^{i\frac{1}{3}\omega} & \mathbf{b}\Sigma e^{-i\frac{1}{3}\omega} \\ \mathbf{b}\Sigma e^{-i\frac{1}{3}\omega} & \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T & \Sigma\mathbf{b}^T e^{i\frac{1}{3}\omega} \\ \Sigma\mathbf{b}^T e^{i\frac{1}{3}} & \mathbf{b}\Sigma e^{-i\frac{1}{3}\omega} & \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T \end{pmatrix},$$

$$\mathbf{g}_4(\omega) = \frac{1}{2\pi}\begin{pmatrix} \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T & \Sigma\mathbf{b}^T e^{i\frac{1}{4}\omega} & \mathbf{0}_d & \mathbf{b}\Sigma e^{-i\frac{1}{4}\omega} \\ \mathbf{b}\Sigma e^{-i\frac{1}{4}\omega} & \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T & \Sigma\mathbf{b}^T e^{i\frac{1}{4}\omega} & \mathbf{0}_d \\ \mathbf{0}_d & \mathbf{b}\Sigma e^{-i\frac{1}{4}\omega} & \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T & \Sigma\mathbf{b}^T e^{i\frac{1}{4}\omega} \\ \Sigma\mathbf{b}^T e^{i\frac{1}{4}\omega} & \mathbf{0}_d & \mathbf{b}\Sigma e^{-i\frac{1}{4}\omega} & \mathbf{\Sigma} + \mathbf{b}\mathbf{\Sigma}\mathbf{b}^T \end{pmatrix},$$

where the $(d \times d)$ zero matrix is denoted by $\mathbf{0}_d$. It can be easily verified that $\mathbf{g}_s(\omega)$ satisfies in all three cases the properties stated in Theorem 3.2.1 (iii) and Corollary 3.2.1, respectively.

Note that the periodically stationary process $(\underline{Y}_t, t \in \mathbb{Z})$ with s periods and predetermined $s \geq 2$ defined in (3.1) becomes a usual stationary linear process if and only if

$$\mathbf{b}_k^{(j)} = \mathbf{b}_k \quad \text{and} \quad \mathbf{\Sigma}_j = \mathbf{\Sigma} \tag{3.9}$$

for all $j = 1, \ldots, s$ and all $k \in \mathbb{Z}$. However, Example 3.2.2 points out that (3.9) is *not* equivalent to a covariance stationary process $(\underline{Y}_t, t \in \mathbb{Z})$.

Example 3.2.2 (The relation of (3.9) and covariance stationarity of $(\underline{Y}_t, t \in \mathbb{Z})$)**.** *Let $(Y_t, t \in \mathbb{Z})$ be a univariate periodically stationary moving average process of order two with $s = 2$ periods. This is a special case of (3.1) with $b_k^{(j)} = 0$ for all $j = 1, 2$ and $k \in \mathbb{Z}\backslash\{0, 1, 2\}$. The corresponding process is called $PMA(2)$ process with two periods, briefly. Now, consider the concrete situation*

$$b_1^{(1)} = 1, \quad b_2^{(1)} = 2, \quad \Sigma_1 = 1,$$
$$b_1^{(2)} = -2, \quad b_2^{(2)} = \frac{1}{2}, \quad \Sigma_2 = 4.$$

Using relation (3.9), the process $(Y_t, t \in \mathbb{Z})$ is not a usual linear process, but computing its generally periodic autocovariance function $\mathbf{\Gamma}_Y(h, m)$ defined in (3.2) results in

$$\Gamma_Y(h, 1) = \Gamma_Y(h, 2) = \Gamma_Y(h)$$

for all $h \in \mathbb{Z}$, that is, $(Y_t, t \in \mathbb{Z})$ is covariance stationary, anyway. Actually, regarding their autocovariance structure, $(Y_t, t \in \mathbb{Z})$ and the $MA(2)$ process $(\widetilde{Y}_t, t \in \mathbb{Z})$ with

$$\widetilde{Y}_t = \widetilde{\epsilon}_t + \widetilde{b}_2\widetilde{\epsilon}_{t-2}, \quad t \in \mathbb{Z}$$

and i.i.d. white noise $\widetilde{\epsilon}_t \sim (0, \frac{2}{\widetilde{b}_2})$ *where* $\widetilde{b}_2 \in \{\frac{\sqrt{65}+9}{4}, \frac{-\sqrt{65}+9}{4}\}$ *are indistinguishable. Due to relation* (3.3), *this property holds also for the corresponding bivariate processes* $(\underline{X}_t, t \in \mathbb{Z})$ *and* $(\underline{\widetilde{X}}_t, t \in \mathbb{Z})$ *and this causes their spectral densities* $\mathbf{f}(\omega)$ *and* $\widetilde{\mathbf{f}}(\omega)$ *to be identical.*

3.3 The test statistic and asymptotic results

3.3.1 Construction of the test statistic

Let $s \geq 2$. Suppose we have d-dimensional data $\underline{Y}_1, \ldots, \underline{Y}_N$ with $N = sM$, $M \in \mathbb{N}$ generated by (3.1) at hand and we are interested in testing for stationarity of the process $(\underline{Y}_t, t \in \mathbb{Z})$, that is, the null hypothesis of interest is

H_0: $(\underline{Y}_t, t \in \mathbb{Z})$ is covariance stationary

against the alternative

H_1: $(\underline{Y}_t, t \in \mathbb{Z})$ is *not* covariance stationary, but periodically stationary with s periods.

To motivate a test statistic for H_0 against H_1 suppose that the spectral density matrix $\mathbf{f}(\omega)$ of the corresponding sd-variate process $(\underline{X}_t, t \in \mathbb{Z})$ is known for all frequencies ω. Hence, we are able to compute $\mathbf{g}(\omega)$ as well, because $\mathbf{d}(\omega)$ is deterministic and known. Under H_0, equation (3.8) in Corollary 3.2.1 holds true and is equivalent to

$$\|\mathbf{G}_{mn}(\omega) - \frac{1}{s}\sum_{j=0}^{s-1} \mathbf{G}_{m+j,n+j}(\omega)\|^2 = 0 \tag{3.10}$$

for all $\omega \in [-\pi, \pi]$ and all $m, n = 1, \ldots, s$, where for a matrix A, $\|A\|$ denotes its Euclidean matrix norm. The previous equation suggests that a way to test the null hypothesis is to test for the squared and normed expression in (3.10) to be equal to zero on the whole interval $\omega \in [-\pi, \pi]$ and for all m, n. Equivalently, integrating and summing-up (3.10) gives

$$\int_{-\pi}^{\pi} \sum_{m,n=1}^{s} \|\mathbf{G}_{mn}(\omega) - \frac{1}{s}\sum_{j=0}^{s-1} \mathbf{G}_{m+j,n+j}(\omega)\|^2 d\omega = 0. \tag{3.11}$$

Because $\mathbf{g}(\omega)$ is unknown in general, let $\widehat{\mathbf{G}}_{mn}(\omega)$ be the canonical nonparametric estimate for $\mathbf{G}_{mn}(\omega)$ obtained by smoothing the so-called *modified periodogram matrix*

$$\widehat{\mathbf{I}}_M(\omega) = \mathbf{d}(\omega)\mathbf{I}_M(\omega)\mathbf{d}^H(\omega), \tag{3.12}$$

where $\mathbf{I}_M(\omega) = \underline{J}_M(\omega)\underline{J}_M^H(\omega)$ is the usual periodogram matrix based on $\underline{X}_1, \ldots, \underline{X}_M$,

$$\underline{J}_M(\omega) = \frac{1}{\sqrt{2\pi M}} \sum_{t=1}^{M} \underline{X}_t e^{-it\omega}$$

denotes the multiple discrete Fourier transform and $\mathbf{d}(\omega)$ is defined previous to equation (3.6) in Section 3.2. Precisely, define

$$\begin{aligned} \widehat{\mathbf{g}}(\omega) &= \frac{1}{M} \sum_{k=-\lfloor \frac{M-1}{2} \rfloor}^{\lfloor \frac{M}{2} \rfloor} K_h(\omega - \omega_k)\widehat{\mathbf{I}}_M(\omega_k) \qquad (3.13) \\ &= \begin{pmatrix} \widehat{\mathbf{G}}_{mn}(\omega) \\ m,n = 1,\ldots,s \end{pmatrix} = \begin{pmatrix} \widehat{g}_{mn}(\omega) \\ m,n = 1,\ldots,sd \end{pmatrix}, \end{aligned}$$

where $\lfloor x \rfloor$ is the integer part of $x \in \mathbb{R}$ and K is a nonnegative symmetric kernel function with compact support $[-\pi, \pi]$ satisfying $\frac{1}{2\pi}\int_{-\pi}^{\pi} K(x)dx = 1$, h is the bandwidth and $K_h(\cdot) = \frac{1}{h}K(\frac{\cdot}{h})$. Recall that N d-dimensional observations correspond to an sd-dimensional realization of length M. Because inference is based on properties of the sd-variate stationary process, the test statistic below is considered to depend on M instead of N.

Now, replacing all unknown quantities in (3.11) by their kernel estimates as derived in (3.13), defining $\widetilde{\mathbf{G}}_{mn}(\omega) = \frac{1}{s}\sum_{j=0}^{s-1} \widehat{\mathbf{G}}_{m+j,n+j}(\omega)$ and introducing the notation

$$\widetilde{\mathbf{g}}(\omega) = \begin{pmatrix} \widetilde{\mathbf{G}}_{mn}(\omega) \\ m,n = 1,\ldots,s \end{pmatrix} = \begin{pmatrix} \widetilde{g}_{mn}(\omega) \\ m,n = 1,\ldots,sd \end{pmatrix}, \qquad (3.14)$$

results in an L_2-type test statistic

$$S_M = Mh^{\frac{1}{2}} \int_{-\pi}^{\pi} \|\widehat{\mathbf{g}}(\omega) - \widetilde{\mathbf{g}}(\omega)\|^2 d\omega, \qquad (3.15)$$

where the inflation factor $Mh^{\frac{1}{2}}$ applied above is reasoned by the variance of S_M as shown in the proof of Theorem 3.3.1. The $(d \times d)$ block entries $\widetilde{\mathbf{G}}_{mn}(\omega)$ of $\widetilde{\mathbf{g}}(\omega)$ are average values computed over all blocks of $\widehat{\mathbf{g}}(\omega)$ that estimate equal quantities under the null hypothesis.

For computational reasons, it is sometimes preferred to avoid matrix norms in the representation of S_M in (3.15). Hence, note that $\widetilde{g}_{mn}(\omega) = \frac{1}{s}\sum_{j=0}^{s-1} \widehat{g}_{m+dj,n+dj}(\omega)$ and

$$S_M = Mh^{\frac{1}{2}} \int_{-\pi}^{\pi} \sum_{m,n=1}^{sd} |\widehat{g}_{mn}(\omega) - \widetilde{g}_{mn}(\omega)|^2 d\omega. \qquad (3.16)$$

At this point, it is worth noting that ratio-type statistics as used in Paparoditis (2000, 2005) for goodness-of-fit testing are not suitable in this situation. This is because we deal exclusively with nonparametric kernel spectral density estimation compared to a combined parametric and nonparametric estimation scheme in the goodness-of-fit setup.

3.3.2 Assumptions and asymptotic results

In order to apply the test statistic S_M, we need its distribution under the null hypothesis. To derive the asymptotic limit of this distribution the following assumptions are imposed.

(A) The d-variate process $(\underline{Y}_t, t \in \mathbb{Z})$ is generated according to (3.1) and comments below. Additionally, $(\underline{\epsilon}_{j+sT}, T \in \mathbb{Z})$ is a d-variate sequence of independent and identically distributed random variables for all $j = 1, \dots, s$. Moreover, $(\underline{\epsilon}_t, t \in \mathbb{Z})$ is assumed to have finite absolute moments of order $16 + \delta$ for some $\delta > 0$, $\sum_{k=-\infty}^{\infty} |k||B_{k;m,n}| < \infty$ for all $m, n = 1, \dots, s$ and the sd-variate process $(\underline{X}_t, t \in \mathbb{Z})$ is supposed to be absolutely regular.

(K) The function K denotes a nonnegative, bounded and Lipschitz continuous kernel with compact support $[-\pi, \pi]$ satisfying $\frac{1}{2\pi}\int_{-\pi}^{\pi} K(u)du = 1$.

(B) The bandwidth h satisfies $h \to 0$ and $Mh \to \infty$ as $M \to \infty$.

Rather than considering only second order moments, we find it more convenient to impose the i.i.d. assumption in (A) that causes the independent sd-variate white noise process $(\underline{e}_t, t \in \mathbb{Z})$ introduced in (3.3) to be identically distributed also. Hence, $(\underline{X}_t, t \in \mathbb{Z})$ becomes a linear process as usually defined. The summability assumption in (A) may be expressed also using the notation in (3.1), but it seems more natural to pose assumptions on the second-order stationary process $(\underline{X}_t, t \in \mathbb{Z})$, because statistical inference is based on its properties.

A unifying contribution on the topic of testing nonparametric and semiparametric hypothesis in the frequency domain is the paper of Eichler (2008), where a general asymptotic theory under certain mixing conditions has been developed. Since this chapter deals with linear processes, asymptotic normality of the proposed test statistic S_M is proved under a set of conditions different to the assumptions used in Eichler (2008).

The following theorem deals with the asymptotic normal distribution of test statistic S_M under the null hypothesis of process $(\underline{Y}_t, t \in \mathbb{Z})$ defined in (3.1) to be covariance stationary.

Theorem 3.3.1 (Asymptotic null distribution of S_M).
Suppose that assumptions (A), (K) and (B) are fulfilled. If H_0 is true, then it holds

$$S_M - \mu_h(K) \xrightarrow{\mathcal{D}} \mathcal{N}(0, \tau^2(K))$$

as $M \to \infty$, where

$$\mu_h(K) = h^{-\frac{1}{2}} \frac{1}{2\pi} \int_{-\pi}^{\pi} K^2(v)dv \int_{-\pi}^{\pi} \Bigg(s(s-1) \left| tr(\mathbf{F}_{11}(\omega)) \right|^2 - 2s \sum_{t=1}^{\lfloor \frac{s-1}{2} \rfloor} \left| tr(\mathbf{F}_{1,1+t}(\omega)) \right|^2 - s \left| tr(\mathbf{F}_{1,1+\frac{s}{2}}(\omega)) \right|^2 1(s \ even) \Bigg) d\omega \tag{3.17}$$

and

$$\tau^2(K) \tag{3.18}$$

$$= \frac{1}{\pi^2} \int_{-2\pi}^{2\pi} \left(\int_{-\pi}^{\pi} K(x)K(x+z)dx \right)^2 dz \int_{-\pi}^{\pi} \Bigg(s(s-1) \left| \sum_{n=1}^{s} tr\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n}^H(\omega) \right) \right|^2 - 2s \sum_{t=1}^{\lfloor \frac{s-1}{2} \rfloor} \left| \sum_{n=1}^{s} tr\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n+t}^H(\omega) \right) \right|^2 - s \left| \sum_{n=1}^{s} tr\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n+\frac{s}{2}}^H(\omega) \right) \right|^2 1(s \ even) \Bigg) d\omega.$$

Here and throughout the chapter, $\xrightarrow{\mathcal{D}}$ denotes convergence in distribution, $tr(\mathbf{A}) = \sum_{k=1}^{d} a_{kk}$ is the trace of a $(d \times d)$ matrix $\mathbf{A}$ and $1(s \ even) = 1$ if $\frac{s}{2} \in \mathbb{N}$ and $1(s \ even) = 0$ otherwise.

Based on Theorem 3.3.1 and for $\alpha \in (0,1)$ a test for the null hypothesis H_0 against the alternative H_1 of asymptotic level α is obtained by rejecting H_0 if

$$\frac{S_M - \mu_h(K)}{\tau(K)} \geq u_{1-\alpha}, \tag{3.19}$$

where $\tau(K) = \sqrt{\tau^2(K)}$ and $u_{1-\alpha}$ is the $(1-\alpha)$-quantile of the standard normal distribution.

Note, $\mu_h(K) = O(h^{-\frac{1}{2}})$, that is, the centralizing term $\mu_h(K)$ tends to infinity with increasing sample size. It is well-known that this property is fulfilled for L_2-type statistics in general [cf. Härdle and Mammen (1993), Paparoditis (2000)]. Moreover, it is worth noting that the test statistic S_M has to be evaluated using the (estimated) modified spectral density $\mathbf{g}(\omega)$, but its asymptotic normal distribution may be expressed using exclusively the usual spectral density $\mathbf{f}(\omega)$.

The following example illustrates how the complicated structure of $\mu_h(K)$ and $\tau^2(K)$ derived in Theorem 3.3.1 simplifies in a special case.

Example 3.3.1 ($\mu_h(K)$ and $\tau^2(K)$ for $d=1$ and $s=2$).
Let $d = 1$ and $s = 2$. Under the assumptions of Theorem 3.3.1, it holds

$$\mu_h(K) = h^{-\frac{1}{2}} \frac{1}{2\pi} \int_{-\pi}^{\pi} K^2(v)dv \int_{-\pi}^{\pi} 2|f_{11}(\omega)|^2 \left(1 - C_{12}(\omega)\right) d\omega$$

and

$$\tau^2(K) = \frac{1}{\pi^2}\int_{-2\pi}^{2\pi}\left(\int_{-\pi}^{\pi} K(x)K(x+z)dx\right)^2 dz \int_{-\pi}^{\pi} 2|f_{11}(\omega)|^4 \left(1 - C_{12}(\omega)\right)^2 d\omega,$$

where $C_{jk}(\omega) = \frac{|f_{jk}(\omega)|^2}{f_{jj}(\omega)f_{kk}(\omega)}$ *is the so-called squared coherence between the two components* j *and* k *of* $(\underline{X}_t, t \in \mathbb{Z})$ *[cf. Hannan (1970)].*

A computationally more attractive version of S_M is given by its discretization

$$\widehat{S}_M = 2\pi h^{\frac{1}{2}} \sum_{j=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor} \|\widehat{\mathbf{g}}(\omega_j) - \widetilde{\mathbf{g}}(\omega_j)\|^2,$$

which is obtained by approximating the integral in (3.15) by its corresponding Riemann sum. This discretized version is asymptotically equivalent to S_M, but its asymptotic distribution derived under H_0 in Theorem 3.3.1 still depends through $\mu_h(K)$ and $\tau^2(K)$ on the unknown quantities $\mathbf{F}_{mn}(\omega)$. By default, these may be estimated nonparametrically and replacing them by their estimates and approximating all unknown integrals in (3.17) and (3.18) by their Riemann sums does not affect the asymptotic distribution of S_M either. These considerations result in the following direct Corollary 3.3.1.

Corollary 3.3.1 (Asymptotic null distribution of $\widehat{S}_M$).
Suppose the assumptions of Theorem 3.3.1 are fulfilled and H_0 *is true. Then it holds*

$$\frac{\widehat{S}_M - \widehat{\mu}_h(K)}{\widehat{\tau}(K)} \xrightarrow{\mathcal{D}} \mathcal{N}(0,1)$$

as $M \to \infty$*, where*

$$\begin{aligned}\widehat{\mu}_h(K) \quad = \quad & h^{-\frac{1}{2}}\frac{1}{2\pi}\int_{-\pi}^{\pi} K^2(v)dv \frac{2\pi}{M}\sum_{j=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor}\Bigg(s(s-1)\left|tr(\widehat{\mathbf{F}}_{11}(\omega))\right|^2 \\ & -2s\sum_{t=1}^{\lfloor\frac{s-1}{2}\rfloor}\left|tr(\widehat{\mathbf{F}}_{1,1+t}(\omega))\right|^2 - s\left|tr(\widehat{\mathbf{F}}_{1,1+\frac{s}{2}}(\omega))\right|^2 1(s\ even)\Bigg)\end{aligned}$$

and

$$\begin{aligned}\widehat{\tau}^2(K) &= \frac{1}{\pi^2}\int_{-2\pi}^{2\pi}\left(\int_{-\pi}^{\pi} K(x)K(x+z)dx\right)^2 dz \\ &\quad\times\frac{2\pi}{M}\sum_{j=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor}\left(s(s-1)\left|\sum_{n=1}^{s} tr\left(\widehat{\mathbf{F}}_{1,n}(\omega_j)\widehat{\mathbf{F}}_{1,n}^H(\omega_j)\right)\right|^2\right. \\ &\quad -2s\sum_{t=1}^{\lfloor\frac{s-1}{2}\rfloor}\left|\sum_{n=1}^{s} tr\left(\widehat{\mathbf{F}}_{1,n}(\omega_j)\widehat{\mathbf{F}}_{1,n+t}^H(\omega_j)\right)\right|^2 \\ &\quad \left. -s\left|\sum_{n=1}^{s} tr\left(\widehat{\mathbf{F}}_{1,n}(\omega_j)\widehat{\mathbf{F}}_{1,n+\frac{s}{2}}^H(\omega_j)\right)\right|^2 1(s\ even)\right).\end{aligned}$$

Therefore, we reject the null hypothesis H_0 if

$$\frac{\widehat{S}_M - \widehat{\mu}_h(K)}{\widehat{\tau}(K)} \geq u_{1-\alpha}, \tag{3.20}$$

which seems to be of more practical relevance compared to (3.19) due to computational convenience. Actually, because of Theorem 3.2.1 (iiib), the representations of $\mu_h(K)$ and $\tau^2(K)$ in (3.17) and (3.18) (and of its approximations $\widehat{\mu}_h(K)$ and $\widehat{\tau}^2(K)$, as well) contain some redundancy, but for notational reasons it seems more convenient to allow for some redundancy in this context.

Before Theorem 3.3.2 below characterizes the behaviour of S_M under the alternative, the following remark discusses the issue of estimating $\mu_h(K)$ and $\tau^2(K)$.

Remark 3.3.1 (On computing $\widehat{\mu}_h(K)$ and $\widehat{\tau}^2(K)$)**.**
Observe, to estimate $\mu_h(K)$ and $\tau^2(K)$ nonparametrically, it is not necessary to compute additional quantities. On the one hand, this is because replacing $\mathbf{F}_{mn}(\omega)$ in (3.17) and (3.18) by $\mathbf{G}_{mn}(\omega)$ does not alter $\mu_h(K)$ and $\tau^2(K)$. And, under H_0, $\widetilde{\mathbf{G}}_{mn}(\omega)$ is some kind of a pooled estimate for $\mathbf{G}_{mn}(\omega)$, which has to be computed for the test statistic S_M anyway and uses most information provided by the data, on the other hand.

The following Theorem 3.3.2 proves S_M (and $\widehat{S}_M$) to be an omnibus test that has power against any alternative belonging to H_1.

Theorem 3.3.2 (Omnibus property of S_M)**.**
Let the assumptions of Theorem 3.3.1 be true and assume that H_0 is wrong, that is, $(\underline{Y}_t, t \in \mathbb{Z})$ is not covariance stationary, but periodically stationary with s periods, $s \geq 2$. If $M \to \infty$, then

$$M^{-1}h^{-\frac{1}{2}}S_M \to \int_{-\pi}^{\pi} \|\mathbf{g}(\omega) - \mathbf{g}_1(\omega)\|^2 d\omega$$

in probability, where $\mathbf{g}_1(\omega)$ *denotes the limit of* $\widetilde{\mathbf{g}}(\omega)$ *in probability under the alternative* H_1.

Since under the alternative $\mathbf{g}(\cdot) - \mathbf{g}_1(\cdot) \neq \mathbf{0}$ due to Theorem 3.2.1, continuity of $\mathbf{g}(\cdot) - \mathbf{g}_1(\cdot)$ implies

$$\int_{-\pi}^{\pi} \|\mathbf{g}(\omega) - \mathbf{g}_1(\omega)\|^2 d\omega > 0.$$

Therefore, S_M is an omnibus test that has power against any alternative belonging to H_1.

3.3.3 Testing the null hypothesis of periodic stationarity

In the setup of an underlying periodically stationary model as, for instance, defined in equation (3.1), typically, the test situation H_0 against H_1 is considered exclusively [cf. Vecchia and Ballerini (1991), Ursu and Duchesne (2009)].

However, it is also of considerable relevance to ask whether the process defined in (3.1) is periodically stationary with a certain number of periods $s_0 < s$. For instance, Franses and Paap (2004) have studied the case of quarterly data, that is, $s = 4$ in detail and, in this setup, the canonical question arises whether the true periodicity s_0 is possibly 2 instead of 4. A test procedure restricted to simple hypotheses and finite number of autocovariance lags that is applicable to this situation is proposed in Lenart, Leśkow and Synowiecki (2008).

Now, suppose $s \geq 2$ in (3.1) and assume that the periodicity of interest s_0 satisfies $s > s_0 \geq 2$ and, additionally, $s' = \frac{s}{s_0} \in \mathbb{N}$. Consider the test situation

$H_0^{(s_0)}$: $(\underline{Y}_t, t \in \mathbb{Z})$ is periodically stationary with s_0 periods

against the alternative

$H_1^{(s_0)}$: $(\underline{Y}_t, t \in \mathbb{Z})$ is periodically stationary with s periods, but *not* with s_0 periods.

To transfer the theory developed in Section 3.3.2 above to test $H_0^{(s_0)}$ against $H_1^{(s_0)}$, observe that $(\underline{Y}_t, t \in \mathbb{Z})$ in (3.1) is periodically stationary with s_0 periods if and only if the d'-dimensional linear process $(\underline{Y}'_t, t \in \mathbb{Z})$ with $d' = s_0 d$ and

$$\underline{Y}'_{j+s'T} = \begin{pmatrix} \underline{Y}_{1+(j-1)s_0+sT} \\ \underline{Y}_{2+(j-1)s_0+sT} \\ \vdots \\ \underline{Y}_{s_0+(j-1)s_0+sT} \end{pmatrix}, \quad j = 1, \ldots, s', \quad T \in \mathbb{Z}, \tag{3.21}$$

is covariance stationary. Therefore, the test situation above may be reformulated and is equivalent to

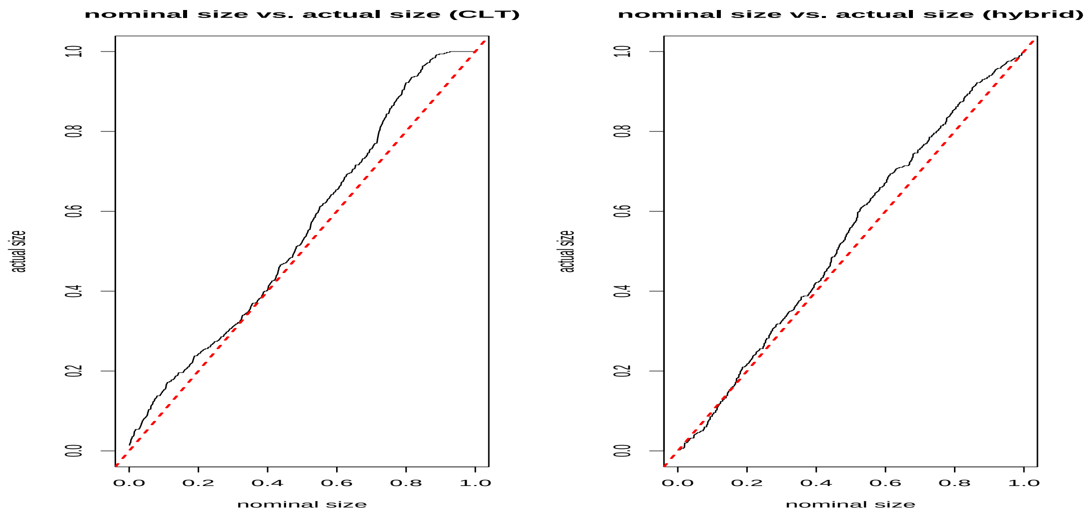

Figure 3.2: P-value plots for Model I using critical values from the CLT (left panel) and from the hybrid bootstrap (right panel) with target indicated by the dashed line.

$H_0^{(s_0)}$: $(\underline{Y}'_t, t \in \mathbb{Z})$ is covariance stationary

against the alternative

$H_1^{(s_0)}$: $(\underline{Y}'_t, t \in \mathbb{Z})$ is *not* covariance stationary, but periodically stationary with s' periods.

Hence, testing $H_0^{(s_0)}$ against $H_1^{(s_0)}$ fits into the setup of Section 3.3.1 and the test statistic S_M derived originally in Section 3.3.2 for testing H_0 against H_1 may be used in this situation as well. Observe, this represents an extension of the previous literature that is typically concerned with testing for stationarity in periodic stationary models or that is restricted to simple hypothesis and finite number of lags of the autocovariance function.

3.4 Simulation studies

In this section we illustrate the performance of the test statistic proposed in Section 3.3 by means of simulations. Suppose we want to test the null hypothesis H_0 of an underlying stationarity process $(\underline{Y}_t, t \in \mathbb{Z})$ against the alternative H_1 of the process being periodically stationary with period $s = 2$. To investigate the behaviour of the test statistic under the null hypothesis, we consider univariate realizations $(d = 1)$ of length $N = 100$ from an $MA(1)$ model

$$Y_t = 0.5\epsilon_{t-1} + \epsilon_t, \quad t \in \mathbb{Z}$$

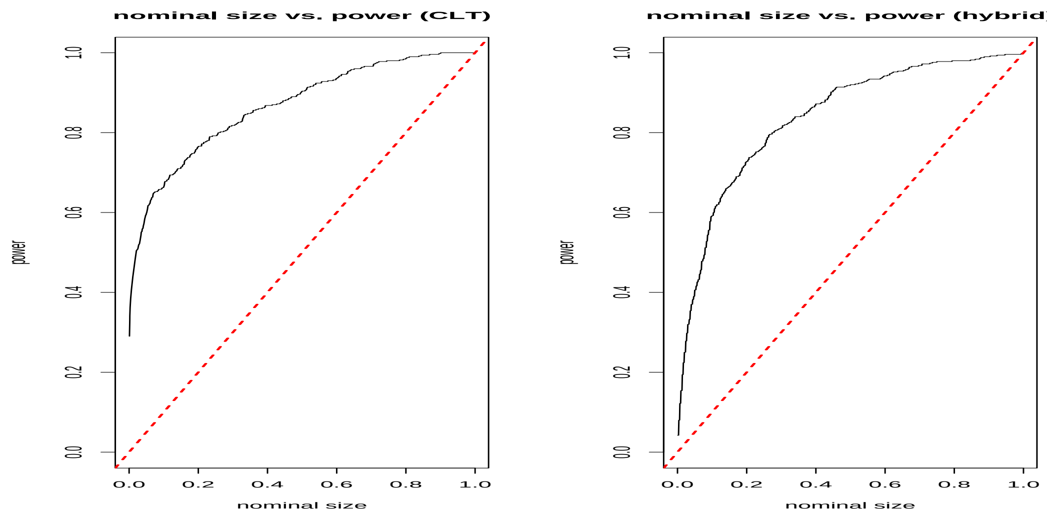

Figure 3.3: Size power curves for Model IIa using critical values from the CLT (left panel) and from the hybrid bootstrap (right panel).

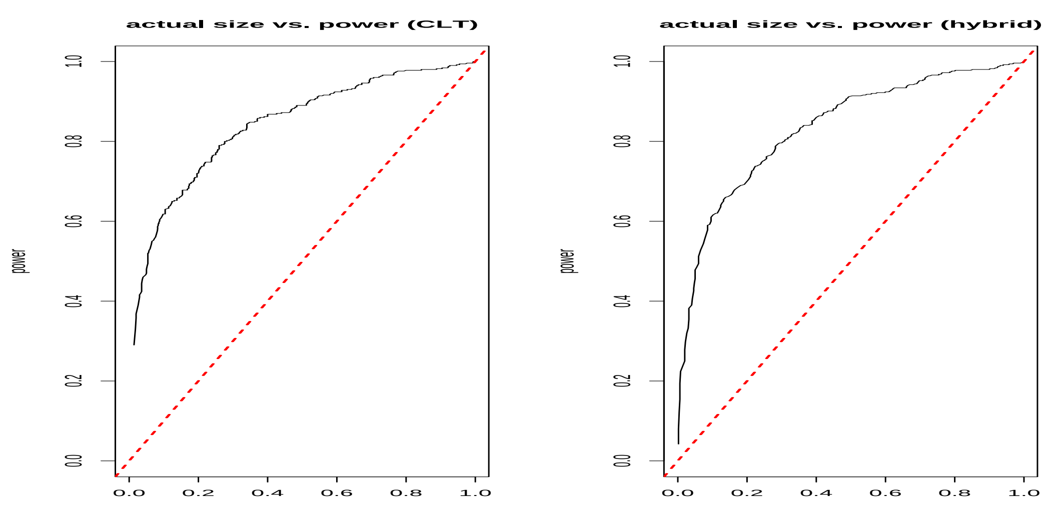

Figure 3.4: Modified size power curves for Model IIa using *actual size* of the test instead of nominal size. Its relation is illustrated in Figure 3.2. Critical values from the CLT (left panel) and from the hybrid bootstrap (right panel) have been used.

with i.i.d. $\epsilon_t \sim \mathcal{N}(0,1)$, which we call Model I in the following.

To analyze the power of the test procedure under the alternative, we consider observations from two periodic $MA(1)$ models

$$\begin{aligned} Y_{1+2T} &= 0.3\epsilon_{1+2T-1} + \epsilon_{1+2T}, \\ Y_{2+2T} &= 0.7\epsilon_{2+2T-1} + \epsilon_{2+2T} \end{aligned}$$

with i.i.d. $\epsilon_t \sim \mathcal{N}(0,1)$ which we call Model IIa and

$$\begin{aligned} Y_{1+2T} &= 0.5\epsilon_{1+2T-1} + \epsilon_{1+2T}, \\ Y_{2+2T} &= 0.5\epsilon_{2+2T-1} + \epsilon_{2+2T} \end{aligned}$$

with independently and normally distributed white noise $(\epsilon_t, t \in \mathbb{Z})$ such that $Var(\epsilon_{j+2T}) = \sigma_j^2$, $j = 1,2$ with $\sigma_1 = 0.8$ and $\sigma_2 = 1.2$ which is indicated as Model IIb. Realizations of Model I, IIa and IIb are shown in Figure 3.1. With $d = 1$, observe that Model I is a special case of Example 3.2.1 and Example 3.3.1, where the modified spectral density $\mathbf{g}_2(\omega)$ of the corresponding bivariate stationary process and the asymptotic distribution of the test statistic S_M under H_0 are given, respectively.

Note that $N = 100$ univariate observations yield to bivariate time series data of length $M = 50$ which is used to estimate the modified spectral density matrix $\mathbf{g}(\omega)$ by $\widehat{\mathbf{g}}(\omega)$ and to compute the related quantity $\widetilde{\mathbf{g}}(\omega)$ for evaluation of the test statistic S_M. In doing so, we have chosen the bandwidth $h = 0.3$ and the Bartlett-Priestley kernel has been used in all simulations.

Among others, Paparoditis (2000, 2005) pointed out that weak convergence of L_2-type statistics of this kind to asymptotic normal distributions is very slow in general. In particular, this holds true for the CLT's presented in Theorem 3.3.1 and in Corollary 3.3.1, respectively. Therefore, an appropriate bootstrap technique should be used to construct a bootstrap version of the test that hopefully shows more reasonable behaviour in small sample situations. To get a bootstrap test that has much power, we have to mimic its distribution under H_0 even if the alternative H_1 is true. In particular, we have to use a bootstrap that works asymptotically for kernel spectral density estimators when the underlying process belongs to the general class of linear time series.

Recently, Jentsch and Kreiss (2010 and Chapter 2 of this thesis) introduced the hybrid bootstrap as an extension of the autoregressive-aided periodogram bootstrap proposed by Kreiss and Paparoditis (2003). They proved not only validity of their proposal for kernel spectral density estimates for linear processes in general, but they also extended this result to the multivariate case. For this reason, it is possible to apply this bootstrap technique when the underlying processes $(\underline{Y}_t, t \in \mathbb{Z})$ is multivariate, that is $d \geq 2$, also. Moreover, the hybrid bootstrap may be applied

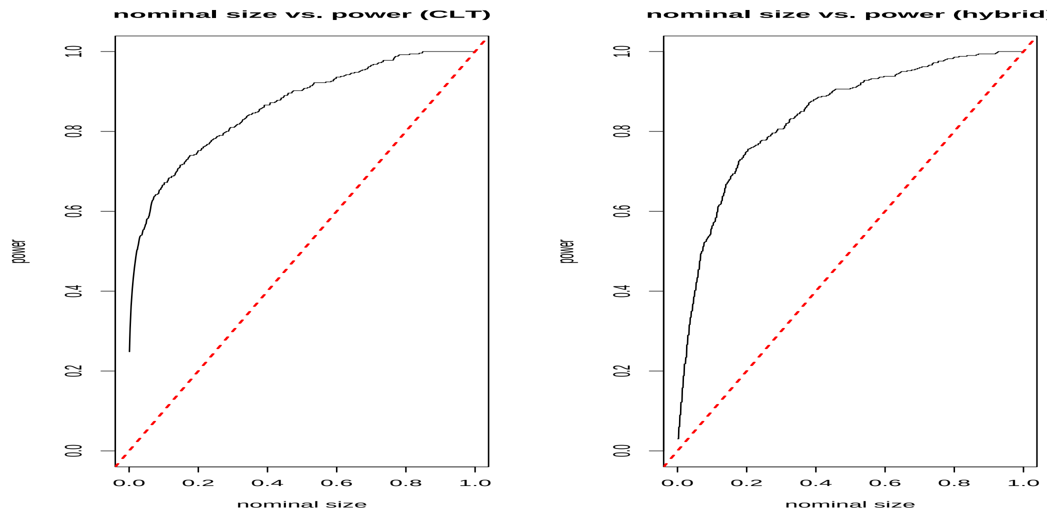

Figure 3.5: Size power curves for Model IIb using critical values from the CLT (left panel) and from the hybrid bootstrap (right panel).

in the situation of Section 3.3.3 as well, where $H_0^{(s_0)}$ against $H_1^{(s_0)}$ is tested. The (univariate) hybrid bootstrap test procedure of level $\alpha \in (0,1)$ may be summarized as follows:

Step 1: Fit an $AR(p)$ model of some (arbitrarily chosen) order $p \in \mathbb{N}$ to the data $Y_1, \ldots, Y_N$ and use a residual bootstrap to get bootstrap replicates $Y_1^+, \ldots, Y_N^+$.

Step 2: Compute the discrete Fourier transform (DFT) $J_N^+(\omega)$ based on $Y_1^+, \ldots, Y_N^+$ and a nonparametric correction term $\widetilde{q}(\omega)$ at the Fourier frequencies $\omega_j = 2\pi \frac{j}{N}$, $j = 1, \ldots, N$.

Step 3: Compute the inverse DFT of the *corrected* DFT $\widetilde{q}(\omega_1)J_N^+(\omega_1), \ldots, \widetilde{q}(\omega_N)J_N^+(\omega_N)$ to obtain bootstrap observations $Y_1^*, \ldots, Y_N^*$ according to

$$Y_t^* = \sqrt{\frac{2\pi}{N}} \sum_{j=1}^{N} \widetilde{q}(\omega_j) J_N^+(\omega_j) e^{it\omega_j}, \quad t = 1, \ldots, N.$$

Step 4: Compute the bootstrap test S_M^* based on $Y_1^*, \ldots, Y_N^*$.

Step 5: Repeat the Steps 1-4 above B times and take the $(1-\alpha)$-quantile of the empirical distribution of $S_{M,1}^*, \ldots, S_{M,B}^*$ to get the α-level bootstrap critical value $c_{M,\alpha}^*$.

Step 5: Finally, reject the null hypothesis if $S_M \geq c_{M,\alpha}^*$.

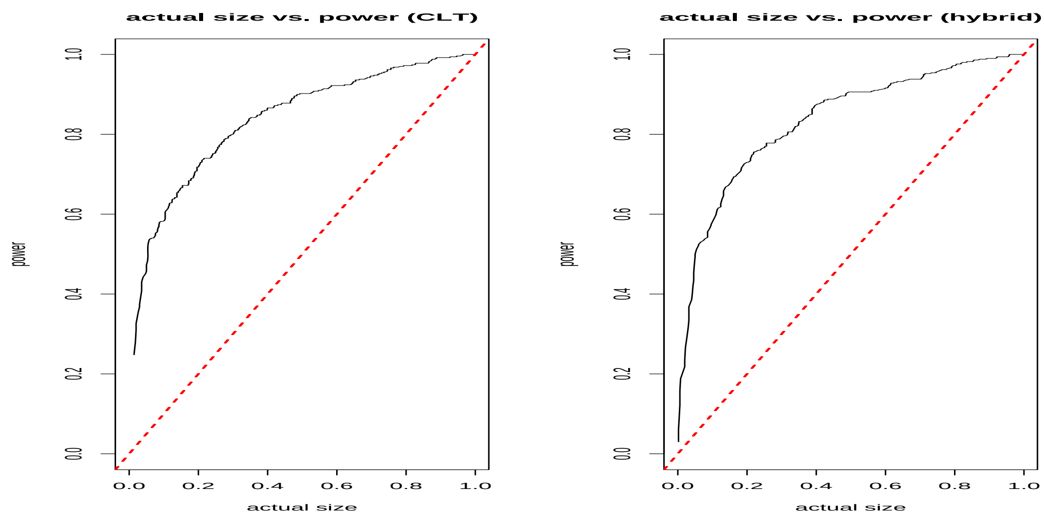

Figure 3.6: Modified size-power curves for Model IIb using *actual size* of the test instead of nominal size. Its relation is illustrated in Figure 3.2. Critical values from the CLT (left panel) and from the hybrid bootstrap (right panel) have been used.

	Model I		
nominal size	0.010	0.050	0.100
actual size (CLT)	0.036	0.086	0.152
actual size (hybrid)	0.006	0.042	0.092

Table 3.1: Comparison of nominal size and actual size for Model I.

	Model IIa			Model IIb		
nominal size	0.010	0.050	0.100	0.010	0.050	0.100
power (CLT)	0.430	0.598	0.664	0.420	0.574	0.666
power (hybrid)	0.154	0.406	0.592	0.122	0.402	0.564

Table 3.2: Comparison of nominal size and power for Model IIa and IIb.

	Model IIa			Model IIb		
actual size	0.010	0.050	0.100	0.010	0.050	0.100
power (CLT)	NA	0.480	0.618	NA	0.474	0.584
power (hybrid)	0.224	0.478	0.616	0.188	0.500	0.584

Table 3.3: Comparison of *actual size* and power for Model IIa and IIb.

The idea behind the hybrid bootstrap is to use a parametric (autoregressive) fit to capture the main dependence features of the data and to apply a nonparametric correction afterwards in the frequency domain to mimic as much as possible the dependence structure in the data. Compare Jentsch and Kreiss (2010 and Chapter 2 of this thesis) for a detailed discussion and the choice of $\widetilde{q}(\omega)$, in particular.

To apply the hybrid bootstrap, some parameters have to be assessed. In our simulation study, we have chosen the autoregressive order $p = 1$ for the residual bootstrap in Step 1 and the smoothing parameter h_b implicitly used for computation of $\widetilde{q}(\omega)$ is chosen to be equal to the bandwidth h for computing S_M, that is $h_b = h = 0.3$, and again the Bartlett-Priestley kernel has been employed. Further, the bootstrap approximation of the distribution of S_M is based on $B = 300$ bootstrap replications.

To examine the behaviour of the proposed test statistic, $T = 500$ data sets have been simulated for all three models under consideration. The results under the null hypothesis of stationarity using Model I are presented in terms of p-value plots in Figure 3.2. The power behaviour of the test under the alternative is illustrated in terms of size power curves in Figure 3.3 for Model IIa and in Figure 3.5 for Model IIb.

In the left panel of Figure 3.2, the p-value plot shows that the test based on critical values from the CLT in Theorem 3.3.1 tends to overreject the null hypothesis systematically. A comparison of both panels in Figure 3.2 demonstrates the gain using critical values from the hybrid bootstrap. In particular for small nominal sizes $\alpha \in [0, 0.2]$ which appear to be crucial for testing purposes, the hybrid bootstrap test in the right panel does not overreject the null hypothesis anymore. For this reason, it seems to be unfair to compare just the usual size-power curves in Figure 3.3 and in Figure 3.5, respectively. Therefore, we present also *modified* size-power curves that use actual sizes instead of nominal sizes on the horizontal axis in Figure 3.4 and in Figure 3.6. Actually, these plots show a very similar shape. The desired effect of the hybrid bootstrap to help the test to hold some predetermined levels is shown in Table 3.1. Also, the relation of nominal (actual) size and power for the test and the bootstrap test are illustrated in Table 3.2 (Table 3.3).

In summary, the test based on critical values obtained from the CLT appears to have good power in both Models IIa and IIb, but tends to overreject the null hypothesis in Model I particularly for small nominal size. As pointed out above, the hybrid bootstrap aided test procedure as shown in the right panel of Figure 3.2 holds the nominal size more accurately under the null and a comparison of modified size-power curves in Figure 3.4 and Figure 3.6, respectively, demonstrates that the bootstrap versions have about the same power as the corresponding tests based on the CLT.

3.5 Proofs

Proof of Theorem 2.1

Let $s \geq 2$. The periodically stationary process $(\underline{Y}_t, t \in \mathbb{Z})$ defined in (3.1) is covariance stationary if and only if its autocovariance function $\mathbf{\Gamma}(h,m)$ introduced in (3.2) does not depend on m, that is,

$$\mathbf{\Gamma}_Y(h,1) = \mathbf{\Gamma}_Y(h,2) = \cdots = \mathbf{\Gamma}_Y(h,s)$$

for all $h \in \mathbb{Z}$. The following equations [cf. for instance Ursu and Duchesne (2009)]

$$\begin{aligned} \mathbf{\Gamma}(h) &= \begin{pmatrix} \mathbf{\Gamma}_{mn}(h) \\ m,n=1,\ldots,s \end{pmatrix} = \begin{pmatrix} \mathbf{\Gamma}_Y(sh+m-n,m) \\ m,n=1,\ldots,s \end{pmatrix}, \qquad (3.22) \\ \mathbf{\Gamma}(h+1) &= \begin{pmatrix} \mathbf{\Gamma}_{mn}(h+1) \\ m,n=1,\ldots,s \end{pmatrix} = \begin{pmatrix} \mathbf{\Gamma}_Y(sh+s+m-n,m) \\ m,n=1,\ldots,s \end{pmatrix} \end{aligned}$$

for $h \in \mathbb{Z}$ establish a relationship between the autocovariance functions $\mathbf{\Gamma}(h)$ and $\mathbf{\Gamma}_Y(h,m)$ of both processes $(\underline{Y}_t, t \in \mathbb{Z})$ and $(\underline{X}_t, t \in \mathbb{Z})$. It can be easily seen that $(\underline{Y}_t, t \in \mathbb{Z})$ is covariance stationary if and only if assertion (ii) of the theorem is satisfied.

To prove the second claimed equivalence, the multivariate inversion formula together with equations (3.6) and (3.22) yields

$$\mathbf{g}(\omega) = \begin{pmatrix} \mathbf{G}_{mn}(\omega) \\ m,n=1,\ldots,s \end{pmatrix} = \frac{1}{2\pi}\sum_{h=-\infty}^{\infty} \begin{pmatrix} \mathbf{\Gamma}_Y(sh+m-n,m)e^{-i\frac{sh+m-n}{s}\omega} \\ m,n=1,\ldots,s \end{pmatrix}. \qquad (3.23)$$

Finally, comparison of coefficients in (3.23) gives that $(\underline{Y}_t, t \in \mathbb{Z})$ is covariance stationary if and only if assertion (iii) is fulfilled, which concludes this proof. □

Proof of Theorem 3.1

To prove the stated central limit theorem for S_M under H_0, it is convenient to deal with its entry-wise representation in (3.16). First, (3.13) and (3.14) yield

$$\widehat{g}_{mn}(\omega) - \widetilde{g}_{mn}(\omega) = \frac{1}{M}\sum_{k=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor} K_h(\omega-\omega_k)\left(-\frac{1}{s}\sum_{j=0}^{s-1}(1-s\delta_{j0})\widehat{I}_{m+dj,n+dj}(\omega_k)\right),$$

where $\delta_{j0} = 1$ if $j = 0$ and $\delta_{j0} = 0$ otherwise. Insertion of the previous identity in (3.16) results in

$$\begin{aligned} &S_M \qquad (3.24) \\ &= Mh^{\frac{1}{2}}\int_{-\pi}^{\pi}\sum_{m,n=1}^{sd}\left|\frac{1}{M}\sum_{k=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor} K_h(\omega-\omega_k)\left(\frac{1}{s}\sum_{j=0}^{s-1}(1-s\delta_{j0})\widehat{I}_{m+dj,n+dj}(\omega_k)\right)\right|^2 d\omega. \end{aligned}$$

The subsequent Lemmas 3.5.1 and 3.5.2 are concerned with the asymptotic behaviour of mean and variance of S_M, respectively. Finally, to complete the proof of Theorem 3.3.1, Lemma 3.5.3 deals with asymptotic normality of S_M.

Lemma 3.5.1 (Computation of $E(S_M)$).
Let $s \geq 2$. Suppose the assumptions (A), (K) and (B) are fulfilled and H_0 is true. Then, it holds

$$\begin{aligned} E(S_M) &= h^{-\frac{1}{2}} \frac{1}{2\pi} \int_{-\pi}^{\pi} K^2(v)dv \int_{-\pi}^{\pi} \Bigg(s(s-1) \left| tr(\mathbf{F}_{11}(\omega)) \right|^2 \\ &\quad -2s \sum_{t=1}^{\lfloor \frac{s-1}{2} \rfloor} \left| tr(\mathbf{F}_{1,1+t}(\omega)) \right|^2 - s \left| tr(\mathbf{F}_{1,1+\frac{s}{2}}(\omega)) \right|^2 1(s \ even) \Bigg) d\omega + o(1), \end{aligned}$$

where the $(d \times d)$ block matrices $\mathbf{F}_{mn}(\omega)$ are defined in (3.5).

Proof.
Expanding the absolute value in (3.24), taking expectation and using the identity

$$E(\widehat{I}_{mn}(\omega_{k_1})\overline{\widehat{I}_{pq}(\omega_{k_2})}) = Cov(\widehat{I}_{mn}(\omega_{k_1}), \widehat{I}_{pq}(\omega_{k_2})) + g_{mn}(\omega_{k_1})\overline{g_{pq}(\omega_{k_2})} + O\left(\frac{1}{M}\right)$$

uniformly in ω_{k_1} and ω_{k_2} results in

$$\begin{aligned} & E(S_M) \\ =\ & Mh^{\frac{1}{2}} \int_{-\pi}^{\pi} \sum_{m,n=1}^{sd} \frac{1}{M^2} \sum_{k_1,k_2=-\lfloor \frac{M-1}{2} \rfloor}^{\lfloor \frac{M}{2} \rfloor} K_h(\omega - \omega_{k_1}) K_h(\omega - \omega_{k_2}) \\ & \frac{1}{s^2} \sum_{j_1,j_2=0}^{s-1} (1 - s\delta_{j_1 0})(1 - s\delta_{j_2 0}) Cov(\widehat{I}_{m+dj_1,n+dj_1}(\omega_{k_1}), \widehat{I}_{m+dj_2,n+dj_2}(\omega_{k_2})) d\omega \\ & + Mh^{\frac{1}{2}} \int_{-\pi}^{\pi} \sum_{m,n=1}^{sd} \frac{1}{M^2} \sum_{k_1,k_2=-\lfloor \frac{M-1}{2} \rfloor}^{\lfloor \frac{M}{2} \rfloor} K_h(\omega - \omega_{k_1}) K_h(\omega - \omega_{k_2}) \\ & \frac{1}{s^2} \sum_{j_1,j_2=0}^{s-1} (1 - s\delta_{j_1 0})(1 - s\delta_{j_2 0}) g_{m+dj_1,n+dj_1}(\omega_{k_1}) \overline{g_{m+dj_2,n+dj_2}(\omega_{k_2})} d\omega + O(h^{\frac{1}{2}}) \\ =\ & A_1 + A_2 + o(1) \end{aligned}$$

with an obvious notation for A_1 and A_2. The second term may be expressed as

$$A_2 = Mh^{\frac{1}{2}} \int_{-\pi}^{\pi} \sum_{m,n=1}^{sd} \left| \frac{1}{M} \sum_{k=-\lfloor \frac{M-1}{2} \rfloor}^{\lfloor \frac{M}{2} \rfloor} K_h(\omega - \omega_k) \frac{1}{s} \sum_{j=0}^{s-1} (1 - s\delta_{j0}) g_{m+dj,n+dj}(\omega_k) \right|^2$$

and vanishes exactly, because of $\sum_{j=0}^{s-1}(1-s\delta_{j0})g_{m+dj,n+dj}(\omega)=0$ for all $\omega\in[-\pi,\pi]$ due to the specific shape of $\mathbf{g}(\omega)$ under the null of covariance stationarity of $(\underline{Y}_t, t\in\mathbb{Z})$ as discussed in Theorem 3.2.1. Now, it remains to consider A_1. Because of $Cov(\widehat{I}_{mn}(\omega_{k_1}),\widehat{I}_{pq}(\omega_{k_2}))=O(\frac{1}{M})$ uniformly in $\omega_{k_1},\omega_{k_2}$ for $|\omega_{k_1}|\neq|\omega_{k_2}|$ and the case $\omega_{k_1}=-\omega_{k_2}$ does not make a contribution asymptotically, it suffices to consider the case $\omega_{k_1}=\omega_{k_2}$ and up to eventually negligible terms, it holds $Cov(\widehat{I}_{mn}(\omega),\widehat{I}_{pq}(\omega))=g_{mp}(\omega)\overline{g_{nq}(\omega)}$, which gives

$$
\begin{aligned}
A_1 &= Mh^{\frac{1}{2}}\int_{-\pi}^{\pi}\sum_{m,n=1}^{sd}\frac{1}{M^2}\sum_{k=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor}(K_h(\omega-\omega_k))^2\\
&\quad\times\frac{1}{s^2}\sum_{j_1,j_2=0}^{s-1}(1-s\delta_{j_10})(1-s\delta_{j_20})g_{m+dj_1,m+dj_2}(\omega_k)\overline{g_{n+dj_1,n+dj_2}(\omega_k)}d\omega+o(1)\\
&= h^{-\frac{1}{2}}\int_{-\pi}^{\pi}\frac{1}{Mh}\sum_{k=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor}K^2\left(\frac{\omega-\omega_k}{h}\right)\\
&\quad\times\frac{1}{s^2}\sum_{j_1,j_2=0}^{s-1}(1-s\delta_{j_10})(1-s\delta_{j_20})\left|\sum_{m=1}^{sd}g_{m+dj_1,m+dj_2}(\omega_k)\right|^2 d\omega+o(1).
\end{aligned}
$$

Approximating the involved Riemann sum by its limiting integral and a standard substitution results in the asymptotically equivalent statistic

$$
h^{-\frac{1}{2}}\frac{1}{2\pi}\int_{-\pi}^{\pi}K^2(v)dv\int_{-\pi}^{\pi}\frac{1}{s^2}\sum_{j_1,j_2=0}^{s-1}(1-s\delta_{j_10})(1-s\delta_{j_20})\left|\sum_{m=1}^{sd}g_{m+dj_1,m+dj_2}(\omega)\right|^2 d\omega.
$$

By manipulating the summation order and, once more, by exploiting the specific block entry-wise structure of $\mathbf{g}(\omega)$ under H_0, the integrand of the second integral above may be expressed as

$$
\begin{aligned}
&\frac{s(s-1)}{s^2}\left|\sum_{m=1}^{sd}g_{mm}(\omega)\right|^2-\frac{2}{s}\sum_{t=1}^{\lfloor\frac{s-1}{2}\rfloor}\left|\sum_{m=1}^{sd}g_{m,m+dt}(\omega)\right|^2\\
&-\frac{1}{s}\left|\sum_{m=1}^{sd}g_{m,m+d\frac{s}{2}}(\omega)\right|^2 1(s\text{ even}).
\end{aligned}\tag{3.25}
$$

Moreover, under H_0, it holds $\sum_{m=1}^{sd}g_{m,m+dt}(\omega)=s\cdot\mathrm{tr}(\mathbf{G}_{1,1+t}(\omega))$. To ease notational issues, this equality together with

$$
\mathbf{G}_{m,n}(\omega)=\mathbf{D}_{m,s}(\omega)\mathbf{F}_{m,n}(\omega)\mathbf{D}_{n,s}^H(\omega)=\mathbf{F}_{m,n}(\omega)e^{-i\frac{m-n}{s}\omega}\tag{3.26}
$$

causes (3.25) to be equal to

$$s(s-1)\left|\mathrm{tr}(\mathbf{F}_{11}(\omega))\right|^2 - 2s\sum_{t=1}^{\lfloor\frac{s-1}{2}\rfloor}\left|\mathrm{tr}(\mathbf{F}_{1,1+t}(\omega))\right|^2 - s\left|\mathrm{tr}(\mathbf{F}_{1,1+\frac{s}{2}}(\omega))\right|^2 1(s \text{ even}),$$

which concludes the proof. □

Lemma 3.5.2 (Computation of $Var(S_M)$)**.**
Let $s \geq 2$. Suppose the assumptions (A), (K) and (B) are fulfilled and H_0 is true. Then, it holds

$$Var(S_M) = \tau^2(K) + o(1),$$

where $\tau^2(K)$ is defined in (3.18).

Proof.
First, consider the second moment $E(S_M^2)$ instead of $Var(S_M)$. Expanding all absolute values in (3.24) gives

$$\begin{aligned} E(S_M^2) \;&=\; M^2 h\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\sum_{m_1,n_1=1}^{sd}\sum_{m_2,n_2=1}^{sd}\frac{1}{M^4}\sum_{k_1,k_2,k_3,k_4=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor} K_h(\omega-\omega_{k_1}) \qquad (3.27)\\ &\times K_h(\omega-\omega_{k_2})K_h(\lambda-\omega_{k_3})K_h(\lambda-\omega_{k_4})\frac{1}{s^4}\sum_{j_1,j_2,j_3,j_4=0}^{s-1}(1-s\delta_{j_10})\\ &\times(1-s\delta_{j_20})(1-s\delta_{j_30})(1-s\delta_{j_40})E\Big[\widehat{I}_{m_1+dj_1,n_1+dj_1}(\omega_{k_1})\\ &\times\overline{\widehat{I}_{m_1+dj_2,n_1+dj_2}(\omega_{k_2})}\widehat{I}_{m_2+dj_3,n_2+dj_3}(\omega_{k_3})\overline{\widehat{I}_{m_2+dj_4,n_2+dj_4}(\omega_{k_4})}\Big]\,d\omega d\lambda. \end{aligned}$$

Due to the identities $\widehat{I}_{m+dj,n+dj}(\omega_k) = I_{m+dj,n+dj}(\omega_k)e^{-i\lfloor\frac{m-n}{d}\rfloor\frac{1}{s}\omega_k}$ and $X_{t,m+dj} = \sum_{\nu=-\infty}^{\infty}\sum_{\mu=1}^{sd}B_{\nu;m+dj,\mu}e_{t-\nu,\mu}$ obtained from (3.3), the expectation on the right side of the last equation above becomes

$$\begin{aligned} &\frac{1}{16\pi^4M^4}\sum_{t_1,\dots,t_8=1}^{M}\sum_{\nu_1,\dots,\nu_8=-\infty}^{\infty}\sum_{\mu_1,\dots,\mu_8=1}^{sd}B_{\nu_1;m_1+dj_1,\mu_1}B_{\nu_2;n_1+dj_1,\mu_2}\\ &\times B_{\nu_3;m_1+dj_2,\mu_3}B_{\nu_4;n_1+dj_2,\mu_4}B_{\nu_5;m_2+dj_3,\mu_5}B_{\nu_6;n_2+dj_3,\mu_6}B_{\nu_7;m_2+dj_4,\mu_7}B_{\nu_8;n_2+dj_4,\mu_8}\\ &\times E\left[e_{t_1-\nu_1,\mu_1}e_{t_2-\nu_2,\mu_2}e_{t_3-\nu_3,\mu_3}e_{t_4-\nu_4,\mu_4}e_{t_5-\nu_5,\mu_5}e_{t_6-\nu_6,\mu_6}e_{t_7-\nu_7,\mu_7}e_{t_8-\nu_8,\mu_8}\right] \qquad (3.28)\\ &\times e^{-i(t_1-t_2)\omega_{k_1}}e^{i(t_3-t_4)\omega_{k_2}}e^{-i(t_5-t_6)\omega_{k_3}}e^{i(t_7-t_8)\omega_{k_4}}\\ &\times e^{-i\lfloor\frac{m_1-n_1}{d}\rfloor\frac{1}{s}\omega_{k_1}}e^{i\lfloor\frac{m_1-n_1}{d}\rfloor\frac{1}{s}\omega_{k_2}}e^{-i\lfloor\frac{m_2-n_2}{d}\rfloor\frac{1}{s}\omega_{k_3}}e^{i\lfloor\frac{m_2-n_2}{d}\rfloor\frac{1}{s}\omega_{k_4}}. \end{aligned}$$

Considering (3.28), in evaluation of (3.27) the cases with largest contribution are those consisting of pairs of e_t's and there are exactly five combinations that do not vanish and make contributions asymptotically to $E(S_M^2)$. The first relevant case is

$$\begin{aligned} t_1-\nu_1 = t_3-\nu_3, \quad t_2-\nu_2 = t_4-\nu_4,\\ t_5-\nu_5 = t_7-\nu_7, \quad t_6-\nu_6 = t_8-\nu_8. \end{aligned}$$

and, due to assumption (A), this combination implies $\omega_{k_1} = \omega_{k_2}$ and $\omega_{k_3} = \omega_{k_4}$ (in the limit). In this situation both integrals with respect to ω and λ in (3.27) separate and its contribution to $E(S_M^2)$ cancels out with $-(E(S_M))^2$ asymptotically when evaluating $Var(S_M)$. All of the other four relevant combinations of index pairs converge to the same limit and as a representative consider

$$t_1 - \nu_1 = t_6 - \nu_6, \quad t_2 - \nu_2 = t_5 - \nu_5,$$
$$t_3 - \nu_3 = t_8 - \nu_8, \quad t_4 - \nu_4 = t_7 - \nu_7.$$

In this setup and due to assumption (A), expression (3.28) is (asymptotically) equal to

$$g_{m_1+dj_1,n_2+dj_3}(\omega_{k_1})\overline{g_{n_1+dj_1,m_2+dj_3}(\omega_{k_1})g_{m_1+dj_2,n_2+dj_4}(\omega_{k_2})}g_{n_1+dj_2,m_2+dj_4}(\omega_{k_2})$$
$$1(\omega_{k_1} = \omega_{k_3})1(\omega_{k_2} = \omega_{k_4}).$$

Now, inserting this term in (3.27), taking all four relevant combinations into considerations which gives a factor 4 and further calculations yield

$$\begin{aligned}
& Var(S_M) \\
= \; & \frac{4h}{M^2} \sum_{k_1,k_2=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor} \left(\int_{-\pi}^{\pi} K_h(\omega - \omega_{k_1})K_h(\omega - \omega_{k_2})d\omega\right)^2 \\
& \times \frac{1}{s^4} \sum_{j_1,j_2,j_3,j_4=0}^{s-1} (1 - s\delta_{j_1 0})(1 - s\delta_{j_2 0})(1 - s\delta_{j_3 0})(1 - s\delta_{j_4 0}) \\
& \times \left| \sum_{m,n=1}^{sd} g_{m+dj_1,n+dj_3}(\omega_{k_1})\overline{g_{m+dj_2,n+dj_4}(\omega_{k_2})}\right|^2 + o(1) \qquad (3.29) \\
= \; & \frac{4h}{M^2} \sum_{k_1,k_2=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor} \left(\int_{-\pi}^{\pi} K_h(\omega - \omega_{k_1})K_h(\omega - \omega_{k_2})d\omega\right)^2 R(\omega_{k_1}, \omega_{k_2}) + o(1)
\end{aligned}$$

with an obvious notation for $R(\omega_{k_1}, \omega_{k_2})$. Approximation of both Riemann sums by their limiting integrals and standard substitutions yield the asymptotically equivalent expression

$$\frac{1}{\pi^2}\int_{-2\pi}^{2\pi}\left(\int_{-\pi}^{\pi} K(x)K(x+z)dx\right)^2 dz \int_{-\pi}^{\pi} R(\omega,\omega)d\omega.$$

To get rid of (most of) the redundancy contained in the sums over j_1, j_2, j_3 and j_4 in $R(\omega) = R(\omega, \omega)$ as defined above in (3.29), consider all nine combinations of the Cartesian product

$$\{\{j_1 = j_3\},\ \{j_1 < j_3\},\ \{j_1 > j_3\}\} \times \{\{j_2 = j_4\},\ \{j_2 < j_4\},\ \{j_2 > j_4\}\}. \qquad (3.30)$$

For instance, using again the specific shape of $\mathbf{g}(\omega)$ under H_0 together with equation (3.26) and $\sum_{m,n=1}^{sd} |g_{mn}(\omega)|^2 = s\sum_{n=1}^{s} \operatorname{tr}\left(\mathbf{G}_{1,n}(\omega)\mathbf{G}_{1,n}^H(\omega)\right)$, the first element $\{j_1 = j_3\} \times \{j_2 = j_4\}$ makes a contribution of

$$\begin{aligned}
&\frac{1}{s^4}\sum_{j_1,j_2=0}^{s-1}(1-s\delta_{j_10})^2(1-s\delta_{j_20})^2\left|\sum_{m,n=1}^{sd} g_{m+dj_1,n+dj_1}(\omega)\overline{g_{m+dj_2,n+dj_2}(\omega)}\right|^2\\
=\ & (s-1)^2\left|\sum_{n=1}^{s}\operatorname{tr}\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n}^H(\omega)\right)\right|^2
\end{aligned}$$

to the asymptotic variance of S_M. Similar results hold for all other combinations in (3.30) and lengthy calculations result in

$$\begin{aligned}
& R(\omega)\\
=\ & (s-1)^2\left|\sum_{n=1}^{s}\operatorname{tr}\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n}^H(\omega)\right)\right|^2 - 4\frac{s-1}{s}\sum_{t=1}^{s-1} t\left|\sum_{n=1}^{s}\operatorname{tr}\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n+t}^H(\omega)\right)\right|^2\\
& +\frac{2}{s}\sum_{t_1,t_2=1}^{s-1} t_1\left|\sum_{n=1}^{s}\operatorname{tr}\left(\mathbf{F}_{1,n+t_1}(\omega)\mathbf{F}_{1,n+t_2}^H(\omega)\right)\right|^2.
\end{aligned}$$

Finally, multiple and tedious manipulations of summation orders and repeated applications of Theorem 3.2.1 give

$$\begin{aligned}
R(\omega) &= s(s-1)\left|\sum_{n=1}^{s}\operatorname{tr}\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n}^H(\omega)\right)\right|^2 - 2s\sum_{t=1}^{\lfloor\frac{s-1}{2}\rfloor}\left|\sum_{n=1}^{s}\operatorname{tr}\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n+t}^H(\omega)\right)\right|^2\\
&\quad -s\left|\sum_{n=1}^{s}\operatorname{tr}\left(\mathbf{F}_{1,n}(\omega)\mathbf{F}_{1,n+\frac{s}{2}}^H(\omega)\right)\right|^2 1(s\text{ even}).
\end{aligned}$$

This completes the proof. $\square$

Lemma 3.5.3 (Asymptotic normality of S_M).
Let $s \geq 2$. Suppose the assumptions (A), (K) and (B) are fulfilled and H_0 is true. Then, it holds that

$$S_M - E(S_M)$$

is asymptotically normally distributed.

Proof.
Due to (3.12), entries of the modified periodogram matrix may be expressed as

$$\widehat{I}_{m+dj,n+dj}(\omega_k) = \left(\frac{1}{2\pi}\sum_{l=-(M-1)}^{M-1}\frac{1}{M}\sum_{t=1}^{M-|l|} X_{t,m+dj}X_{t+|l|,n+dj}e^{-il\omega_k}\right)\left(e^{-i\lfloor\frac{m-n}{d}\rfloor\frac{\omega_k}{s}}\right).$$

Inserting this identity in (3.24) and a change of the summation order yields

$$\begin{aligned} S_M &= Mh^{\frac{1}{2}} \int_{-\pi}^{\pi} \sum_{m,n=1}^{sd} \left| \frac{1}{2\pi} \sum_{l=-(M-1)}^{M-1} \left(\frac{1}{Mh} \sum_{k=-\lfloor \frac{M-1}{2} \rfloor}^{\lfloor \frac{M}{2} \rfloor} K(\frac{\omega - \omega_k}{h}) e^{-il\omega_k} e^{-i\lfloor \frac{m-n}{d} \rfloor \frac{\omega_k}{s}} \right) \right. \\ & \quad \times \left. \left(-\frac{1}{s} \sum_{j=0}^{s-1} (1 - s\delta_{j0}) \frac{1}{M} \sum_{t=1}^{M-|l|} X_{t,m+dj} X_{t+|l|,n+dj} \right) \right|^2 d\omega. \end{aligned}$$

The asymptotically equivalent statistic

$$\begin{aligned} & Mh^{\frac{1}{2}} \int_{-\pi}^{\pi} \sum_{m,n=1}^{sd} \left| \frac{1}{2\pi} \sum_{l=-(M-1)}^{M-1} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} K(v) e^{-il(\omega - hv)} e^{-i\lfloor \frac{m-n}{d} \rfloor \frac{\omega - hv}{s}} dv \right) \right. \\ & \quad \times \left. \left(-\frac{1}{s} \sum_{j=0}^{s-1} (1 - s\delta_{j0}) \frac{1}{M} \sum_{t=1}^{M-|l|} X_{t,m+dj} X_{t+|l|,n+dj} \right) \right|^2 d\omega \end{aligned}$$

obtained by approximating the involved Riemann sum by its limiting integral and a standard substitution, may be simplified further using $|e^{-i\lfloor \frac{m-n}{d} \rfloor \frac{\omega}{s}}| = 1$, $|e^{i\lfloor \frac{m-n}{d} \rfloor \frac{hv}{s}} - 1| = O(h)$ and defining $k(l,h) = \frac{1}{2\pi} \int_{-\pi}^{\pi} K(v) cos(lhv) dv$ and

$$a_l = k(l,h) \left(-\frac{1}{s} \sum_{j=0}^{s-1} (1 - s\delta_{j0}) \frac{1}{M} \sum_{t=1}^{M-|l|} X_{t,m+dj} X_{t+|l|,n+dj} \right). \tag{3.31}$$

Altogether, this results in the following representation of S_M and an application of Parsevals' identity gives

$$\begin{aligned} S_M &= \frac{1}{\pi} Mh^{\frac{1}{2}} \sum_{m,n=1}^{sd} \frac{1}{\pi} \int_{-\pi}^{\pi} \left| \frac{a_0}{2} + \sum_{l=1}^{M-1} a_l cos(lx) \right|^2 d\omega + o_P(1) \\ &= \frac{1}{\pi} Mh^{\frac{1}{2}} \sum_{m,n=1}^{sd} \left(\frac{a_0^2}{2} + \sum_{l=1}^{M-1} a_l^2 \right) + o_P(1). \end{aligned}$$

Now, insertion of formula (3.31) in the last expression on the right-hand side above,

$a_0^2 = O(\frac{1}{M})$ and some summation manipulations yield

$$\begin{aligned} S_M \;=\;& \frac{h^{\frac{1}{2}}}{\pi M}\frac{1}{s^2}\sum_{j_1,j_2=0}^{s-1}(1-s\delta_{j_10})(1-s\delta_{j_20}) \\ &\times\sum_{t=2}^{M}\left(\sum_{l=1}^{t-1}k^2(l,h)\left(\sum_{m=1}^{sd}X_{t-l,m+dj_1}X_{t-l,m+dj_2}\right)\right)\left(\sum_{n=1}^{sd}X_{t,n+dj_1}X_{t,n+dj_2}\right) \\ &+\frac{h^{\frac{1}{2}}}{\pi M}\frac{1}{s^2}\sum_{j_1,j_2=0}^{s-1}(1-s\delta_{j_10})(1-s\delta_{j_20}) \\ &\times\sum_{\substack{t_1,t_2=2\\ t_1\neq t_2}}^{M}\left(\sum_{l=1}^{min(t_1-1,t_2-1)}k^2(l,h)\left(\sum_{m=1}^{sd}X_{t_1-l,m+dj_1}X_{t_2-l,m+dj_2}\right)\right) \\ &\times\left(\sum_{n=1}^{sd}X_{t_1,n+dj_1}X_{t_2,n+dj_2}\right)+o_P(1). \end{aligned}$$

The first term on the last right-hand side vanishes asymptotically in $S_M - E(S_M)$ with $E(S_M)$ derived in Lemma 3.5.1. Finally, the second term can be treated with Theorem 2.1 in Gao and Hong (2008) and the claimed asymptotic normality follows consequently from the assumptions posed in Assumption (A). □

Proof of Theorem 3.2

Using identity (3.24) and due to convergence in probability of

$$\frac{1}{M}\sum_{k=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor}K_h(\omega-\omega_k)\left(-\frac{1}{s}\sum_{j=0}^{s-1}(1-s\delta_{j0})\widehat{I}_{m+dj,n+dj}(\omega_k)\right)$$

to $-\frac{1}{s}\sum_{j=0}^{s-1}(1-s\delta_{j0})g_{m+dj,n+dj}(\omega)$ uniformly in ω, it holds

$$\begin{aligned} & M^{-1}h^{-\frac{1}{2}}S_M \\ =\;& \int_{-\pi}^{\pi}\sum_{m,n=1}^{sd}\left|\frac{1}{M}\sum_{k=-\lfloor\frac{M-1}{2}\rfloor}^{\lfloor\frac{M}{2}\rfloor}K_h(\omega-\omega_k)\left(-\frac{1}{s}\sum_{j=0}^{s-1}(1-s\delta_{j0})\widehat{I}_{m+dj,n+dj}(\omega_k)\right)\right|^2 d\omega \\ =\;& \int_{-\pi}^{\pi}\sum_{m,n=1}^{sd}\left|-\frac{1}{s}\sum_{j=0}^{s-1}(1-s\delta_{j0})g_{m+dj,n+dj}(\omega)\right|^2 d\omega+o_P(1) \\ =\;& \int_{-\pi}^{\pi}\|\mathbf{g}(\omega)-\mathbf{g}_1(\omega)\|^2 d\omega+o_P(1) \end{aligned}$$

and this completes the proof. □

Bibliography

[1] Bentarzi, M. and Hallin, M. (1994): On the invertibility of periodic moving-average models. *J. Time Ser. Anal.*, 15, No.3, 263-268.

[2] Bentarzi, M. and Hallin, M. (1998): Spectral factorization of periodically correlated MA(1) processes. *J. Appl. Probab.*, 35, No.1, 46-54.

[3] Berkowitz, J. and Diebold, F.X. (1997): Bootstrapping Multivariate Spectra. Preprint.

[4] Bickel, J.P. and Freedman, D.A. (1981): Some Asymptotic Theory for the Bootstrap. *Ann. Statist.*, 9, No. 6, 1196-1217.

[5] Brillinger, D. (1981): *Time Series: Data Analysis and Theory.* Holden-Day, San Francisco.

[6] Brockwell, P.J. and Davis, R.A. (1991): *Time Series: Theory and Methods.* Springer, New York.

[7] Bühlmann, P. (1997): Sieve bootstrap for time series. *Bernoulli*, 3, No. 2, 123-148.

[8] Bühlmann, P. (2002): Bootstraps for Time Series. *Statist. Sci.*, 17, 52-72.

[9] Cipra, T. (1985): Periodic moving average process. *Apl. Mat.*, 30, 218-229.

[10] Dahlhaus, R. (1985): Asymptotic normality of spectral estimates. *J. Multivariate Anal.*, 16, 412-431.

[11] Dahlhaus, R. and Janas, D. (1996): A frequency domain bootstrap for ratio statistics in time series analysis. *Ann. Statist.*, 24, 1934-1963.

[12] Dai, M. and Guo, W. (2004): Multivariate Spectral Analysis Using Cholesky Decomposition. *Biometrika*, 91, No. 3, 629-643.

[13] Efron, B. (1979): Bootstrap Methods: Another Look at the Jackknife. *Ann. Statist.*, 7, 1-26.

[14] Franke, J. and Härdle, W. (1992): On Bootstrapping Kernel Spectral Estimates. *Ann. Statist.*, 20, No. 1, 121-145.

[15] Franses, P.H. and Paap, R. (2004): *Periodic Time Series Models.* Oxford University Press, New York.

[16] Gao, J. and Hong, Y. (2008): Central limit theorems for generalized U-statistics with applications in nonparametric specification. *J. Nonparametric Stat.*, 20, No. 1, 61-76.

[17] Gladyshev, E.G. (1961): Periodically correlated random sequences. *Soviet Mathematics.*, 2, 385-388.

[18] Guo, W. and Dai, M. (2006): Multivariate time-dependent Spectral Analysis Using Cholesky Decomposition. *Stat. Sin.*, 16, No. 3, 825-845.

[19] Hannan, E.J. (1970): *Multiple Time Series.* Wiley, New York.

[20] Härdle, W. and Mammen, E. (1993): Comparing nonparametric versus parametric regression fits. *Ann. Stat.*, 21, No. 4, 1926-1947.

[21] Härdle, W, Horowitz, J. and Kreiss, J.-P. (2003): Bootstrap Methods for Time Series. *Int. Statist. Rev.*, 71, No. 2, 435-459.

[22] Herwartz, H. (1998): Testing periodicity in time series models - A recommendation of bootstrap models. *Comput. Stat.*, 13, No. 3, 283-299.

[23] Hurvich, C.M. and Zeger, S.L. (1987): Frequency domain bootstrap methods for time series. Technical Report 87-115, Graduate School of Business Administration, New York Univ.

[24] Janas, D. and Dahlhaus, R. (1994): A frequency domain bootstrap for time series. In *Computationally Intensive Statistical Methods. Proc. 26th Symposium on the Interface* (J. Stall and A. Lehman, eds.) 423-425. Interface Foundation of North America, Fairfax Station, VA.

[25] Jentsch, C. and Kreiss, J.-P. (2010): The multiple Hybrid Bootstrap - Resampling multivariate linear processes. *J. Multivariate Anal.*, 101, 2320 - 2345.

[26] Jentsch, C. (2010): A new frequency domain approach of testing for covariance stationarity and for periodic stationarity in multivariate linear processes. *J. Time Ser. Anal.* Submitted.

[27] Jones, R.H. and Brelsford, W. (1967): Time series with periodic structure. *Biometrika.*, 54, 403-408.

[28] Kirch, C. and Politis, D. (2009): TFT-Bootstrap: Resampling in the Frequency Domain to obtain Replicates in the Time Domain. Preprint.

[29] Kreiss, J.-P. (1984): *Adaptive Estimation and Testing in ARMA Models - the Multivariate Case.* Unpublished manuscript.

[30] Kreiss, J.-P. (1988): Asymptotic Statistical Inference for a Class of Stochastic Processes. Habilitationsschrift (105 Seiten), Fachbereich Mathematik der Universität Hamburg.

[31] Kreiss, J.-P. and Franke, J. (1992): Bootstrapping Stationary Autoregressive Moving-Average Models. *J. Time Ser. Anal.*, 13, No. 4, 297-317.

[32] Kreiss, J.-P. (1999): *Residual and Wild Bootstrap for Infinite Order Autoregression.* Unpublished manuscript.

[33] Kreiss, J.-P. and Paparoditis, E. (2003): Autoregressive-aided periodogram bootstrap for time series. *Ann. Statist.*, 31, No. 6, 1923-1955.

[34] Kreiss, J.-P. and Neuhaus, G. (2006): *Einführung in die Zeitreihenanalyse.* Springer, Berlin.

[35] Künsch, H.R. (1989): The Jackknife and the Bootstrap for General Stationary Observations. *Ann. Statist.*, 17, No. 3, 1217-1241.

[36] Lahiri, S.N. (2003): *Resampling Methods for Dependent Data.* Springer, New York.

[37] Lenart, Ł., Leśkow, J. and Synowiecki, R. (2008): Subsampling in testing autocovariance for periodically correlated series. *J. Time Ser. Anal.*, 29, No. 6, 995-1018.

[38] Lütkepohl, H. (1996): *Handbook of Matrices.* Wiley, Chichester.

[39] Lütkepohl, H. (2005): *New Introduction to Multiple Time Series Analysis.* Springer, Berlin.

[40] Lund, R., Shao, Q. and Basawa, I. (2006): Parsimonious periodic time series modeling. *Aust. N. Z. J. Stat.*, 48, No. 1, 33-47.

[41] Mallows, C.L. (1972): A Note on Asymptotic Joint Normality. *Ann. Math. Statist.*, 43, 508-515.

[42] McLeod, A.I. (1993): Parsimony, model adequacy, and periodic autocorrelation in time series forecasting. *Int. Stat. Rev.*, 61, No. 3, 387-393.

[43] McLeod, A.I. (1994): Diagnostic checking periodic autoregression models with applications. *J. Time Ser. Anal.*, 15, No. 2, 221-233.

[44] Neumann, M.H. and Paparoditis, E. (2008): Goodness-of-fit Tests for Markovian Time Series Models. *Bernoulli*, 14, No. 1, 14-46.

[45] Pagano, M. (1978): On periodic and multiple autoregressions. *Ann. Stat.*, 6, 1310-1317.

[46] Paparoditis, E. (1996): Bootstrapping Autoregressive and Moving Average Parameter Estimates of Infinite Order Vector Autoregressive Processes. *J. Multivariate Anal.*, 57, 277-296.

[47] Paparoditis, E. and Politis, D.N. (1999): The Local Bootstrap for periodogram statistics. *J. Time Ser. Anal.*, 20, 193-222.

[48] Paparoditis, E. (2000): Spectral density based goodness-of-fit tests for time series models. *Scand. J. Stat.*, 27, No.1, 143-176.

[49] Paparoditis, E. (2002): Frequency Domain Bootstrap for Time Series. In *Empirical Process Techniques for Dependent Data* (H. Dehling, T. Mikosch and M. Sørensen, eds.) 365-381. Birkhäuser, Boston.

[50] Paparoditis, E. (2005): Testing the Fit of a Vector Autoregressive Moving Average Model. *J. Time Ser. Anal.*, 16, No. 4, 543-568.

[51] Priestley, M.B. (1981): *Spectral Analysis and Time Series.* Academic Press, New York.

[52] Rosenblatt, M. (1985): *Stationary Sequences and Random Fields.* Birkhäuser, Boston.

[53] Sakai, H. (1991): On the spectral density matrix of periodic ARMA process. *J. Time Ser. Anal.*, 12, No. 1, 73-82.

[54] Sergides, M. and Paparoditis, E. (2007): Bootstrapping the Local Periodogram of Locally Stationary Processes. *J. Time Ser. Anal.*, 29, No. 2, 264-299.

[55] Shao, X. and Wu, W.B. (2007): Asymptotic spectral theory for nonlinear time series. *Ann. Statist.*, 35, No. 4, 1773-1801.

[56] Tiao, G.C. and Grupe, M.R. (1980): Hidden periodic autoregressive moving average models in time series data. *Biometrika.*, 67, 365-373.

[57] Troutman, B.M. (1979): Some results in periodic autoregression. *Biometrika.*, 66, 219-228.

[58] Ula, A.T. (1990): Periodic covariance stationarity of multivariate periodic autoregressive moving average processes. *Water Resources Research.*, 26, 855-861.

[59] Ula, A.T. (1993): Forecasting of multivariate periodic autoregressive moving-average processes. *J. Time Ser. Anal.*, 14, 645-657.

[60] Ursu, E. and Duchesne, P. (2009): On modelling and diagnostic checking of vector periodic autoregressive time series models. *J. Time Ser. Anal.*, 30, 70-96.

[61] Vecchia, A.V. (1985a): Periodic autoregressive moving-average (PARMA) modeling with appications to water resources. *Water Resources Bulletin*, 21, 721-730.

[62] Vecchia, A.V. (1985b): Maximum likelihood estimation for periodic autoregressive moving average models. *Technometrics*, 27, 375-384.

[63] Vecchia, A.V. and Ballerini, R. (1991): Testing for periodic autocorrelations in seasonal time series data. *Biometrika.*, 78, 53-63.

[64] Whittle, P. (1963): On the Fitting of Multivariate Autoregressions, and the Approximate Canonical Factorization of a Spectral Density Matrix. *Biometrika*, 50, No. 1/2, 129-134.

List of Figures

List of Tables

English summary

In the first part of this thesis, which may be included to the field of time series analysis, a subdiscipline of mathematical statistics, a new bootstrap procedure for dependent data is proposed and its properties are discussed. Under the assumption of an underlying linear time series model, the idea of the *autoregressive aided periodogram bootstrap* (AAPB) of Kreiss and Paparoditis (2003) is reconsidered and in two directions generalized and complemented, respectively. On the one hand, the AAPB is modified in such a way that it is eventually able to generate bootstrap observations in the time domain, which was not possible for the AAPB. On the other hand, not only univariate, but also multivariate processes of arbitrary dimension are considered. It is shown that the here proposed multiple hybrid bootstrap (mHB) that includes the AAPB as a special case, is consistent under quite general assumptions for the sample mean and for kernel spectral density estimates. Furthermore, it is shown that the mHB also works in the univariate case under general assumptions for autocorrelations, but for autocovariances only under additional assumptions. However, in the multivariate case, additional assumptions are required in both situations to obtain asymptotically consistent procedures.

The second part of this thesis deals with multivariate linear periodically stationary models, which generalize the usual stationary linear models in that effect that their coefficients are no longer assumed to be constant over time, but to behave periodically. These models may be represented as higher-dimensional stationary models and it is shown that the autocovariance structure as well as the spectral density of this higher-dimensional process form upon a specific pattern if and only if the underlying process is actually not just periodically stationary, but also stationary in the classical sense. To test for stationarity, a test statistic based on nonparametric spectral density estimates is constructed that takes advantage of this specific shape. The asymptotic normal distribution of the test statistic is derived and it is shown that the test has power asymptotically against any alternative belonging to the class of periodically stationary models. Moreover, it is demonstrated how this test statistic may be used to test for periodic stationarity with shorter period. Because it is well-known that the convergence of test statistics based on nonparametric spectral density estimates is quite slow, the hybrid bootstrap discussed in the first part of this thesis is used to obtain critical values that are more adequate than those from the central limit theorem.

Deutsche Zusammenfassung

Im ersten Teil der vorliegenden Arbeit, die sich der Zeitreihenanalyse, einem Teilgebiet der Mathematischen Statistik, zuordnen lässt, wird ein neues Bootstrapverfahren für abhängige Daten vorgeschlagen und dessen Eigenschaften werden diskutiert. Unter der Annahme eines zugrundeliegenden linearen Zeitreihenmodells, wird die Idee des *autoregressive aided periodogram bootstrap* (AAPB) von Kreiss und Paparoditis (2003) neu aufgegriffen und in zwei Richtungen verallgemeinert bzw. ergänzt. Zum einen wird das dort untersuchte Bootstrapverfahren so modifiziert, dass es schließlich in der Lage ist, Beobachtungen im Zeitbereich zu erzeugen, was dem AAPB nicht möglich war. Zum anderen werden nicht nur univariate, sondern auch multivariate Prozesse beliebiger Dimension betrachtet. Es wird gezeigt, dass der hier vorgeschlagene multiple hybride Bootstrap (mHB), der den AAPB als Spezialfall enthält, unter allgemeinen Voraussetzungen für den Mittelwert und für Spektraldichtekernschätzer konsistent ist. Weiter wird gezeigt, dass der mHB im univariaten Fall ebenso unter allgemeinen Bedingungen für Autokorrelationen funktioniert, für Autokovarianzen jedoch nur unter Zusatzannahmen. Im multivariaten Fall hingegen, benötigt man für beide Situationen Zusatzbedingungen, um asymptotisch konsistente Verfahren zu erhalten.

Der zweite Teil dieser Arbeit beschäftigt sich mit multivariaten linearen periodisch stationären Modellen, welche die üblichen stationären linearen Modelle dahingehend verallgemeinern, dass die Modellparameter nicht mehr konstant über die Zeit sind, sondern sich periodisch verhalten. Diese Modelle lassen sich als höherdimensionale stationäre Modelle auffassen und es wird gezeigt, dass die Autokovarianzstruktur sowie die Gestalt der Spektraldichte dieses höherdimensionalen Prozesses genau dann einem bestimmten Muster folgt, wenn der zugrundeliegende Prozess tatsächlich nicht nur periodisch stationär, sondern auch stationär im klassischen Sinn ist. Zum Testen auf Stationarität wird eine Teststatistik basierend auf nichtparametrischen Spektraldichteschätzern konstruiert, die sich diese besondere Struktur zu Nutze macht. Die asymptotische Verteilung der Teststatistik unter der Hypothese wird in Form eines zentralen Grenzwertsatzes hergeleitet und es wird gezeigt, dass der Test asymptotisch jede feste Alternative in der Klasse der periodisch stationären linearen Modelle erkennt. Ebenso wird demonstriert, wie die Teststatistik auch benutzt werden kann, um die Hypothese periodischer Stationarität mit kürzerer Periode zu testen. Da hinlänglich bekannt ist, dass die Konvergenz von Teststatstiken basierend auf nichtparametrschen Spektraldichteschätzern recht langsam ist, wird in einer Simulationsstudie das hybride Bootstrapverfahren aus dem ersten Teil dieser Arbeit benutzt, um geeignetere kritische Werte als mit dem zentralen Grenzwertsatz zu erhalten.

Lebenslauf

Persönliche Daten

Name	Carsten Jentsch
geboren am	27.12.1979
in	Holzminden

Werdegang

03/2007 - dato	Wissenschaftlicher Mitarbeiter am Institut für Mathematische Stochastik der Technischen Universität Braunschweig
10/2001 - 02/2007	Studium der Mathematik mit Nebenfach Betriebswirtschaftslehre an der Technischen Universität Braunschweig, Abschluss: Diplom in Mathematik
07/2000 - 04/2001	Grundwehrdienst in Braunschweig und Stadtoldendorf
06/2000	Abitur am Campe Gymnasium in Holzminden